

# Diffusive estimates for random walks on stationary random graphs of polynomial growth

Shirshendu Ganguly

James R. Lee

Yuval Peres

## Abstract

Let  $(G, \rho)$  be a stationary random graph, and use  $B_\rho^G(r)$  to denote the ball of radius  $r$  about  $\rho$  in  $G$ . Suppose that  $(G, \rho)$  has annealed polynomial growth, in the sense that  $\mathbb{E}[|B_\rho^G(r)|] \leq O(r^k)$  for some  $k > 0$  and every  $r \geq 1$ .

Then there is an infinite sequence of times  $\{t_n\}$  at which the random walk  $\{X_t\}$  on  $(G, \rho)$  is at most diffusive: Almost surely (over the choice of  $(G, \rho)$ ), there is a number  $C > 0$  such that

$$\mathbb{E} \left[ \text{dist}_G(X_0, X_{t_n})^2 \mid X_0 = \rho, (G, \rho) \right] \leq C t_n \quad \forall n \geq 1.$$

This result is new even in the case when  $G$  is a stationary random subgraph of  $\mathbb{Z}^d$ . Combined with the work of Benjamini, Duminil-Copin, Kozma, and Yadin (2015), it implies that  $G$  almost surely does not admit a non-constant harmonic function of sublinear growth.

To complement this, we argue that passing to a subsequence of times  $\{t_n\}$  is necessary, as there are stationary random graphs of (almost sure) polynomial growth where the random walk is almost surely superdiffusive at an infinite subset of times.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	The absence of non-constant sublinear growth harmonic functions	3
1.2	Speed via Euclidean embeddings	4
1.3	A deterministic example: Planar graphs	6
<b>2</b>	<b>Martingales, embeddings, and growth rates</b>	<b>8</b>
2.1	Control by martingales	8
2.2	Embeddings and growth rates	11
<b>3</b>	<b>Diffusive estimates</b>	<b>13</b>
3.1	The mass transport principle	13
3.1.1	Subexponential growth and reversibility	15
3.2	Choosing a good Følner set	16
3.3	Multi-scale control of growth functionals	18
<b>4</b>	<b>Existence of exceptional times</b>	<b>21</b>
4.1	Trees of graphs	21
4.2	Stretched expanders and the rate of escape	22
4.3	The recursive construction	24
4.4	Convergence to a stationary random graph	26

# 1 Introduction

It is a classical fact that the standard random walk  $\{X_n\}$  on  $\mathbb{Z}^d$  exhibits *diffusive behavior*:  $\mathbb{E} \|X_0 - X_n\|_2^2 \asymp n$ . The well-known estimates of Varopoulos and Carne [Car85, Var85] show that, if  $\{X_n\}$  is random walk on a graph  $G$  of polynomial growth, then the speed can be at most slightly superdiffusive:  $\mathbb{E} d_G(X_0, X_n)^2 \leq O(n \log n)$ , where  $d_G$  is the graph metric on  $G$ .

Kesten [Kes86] examined the distribution of the random walk on percolation clusters in  $\mathbb{Z}^d$ . Suppose that  $(G, \rho)$  is a stationary random subgraph of  $\mathbb{Z}^d$ . This means that if  $\{X_n\}$  is the random walk conditioned on  $G$  with  $X_0 = \rho$ , then  $(G, X_0) \stackrel{\text{law}}{=} (G, X_1)$ . Kesten's argument can be used to show that, in this case,

$$\forall n \geq 1, \mathbb{E} \|X_0 - X_n\|_2^2 \leq O(n) \quad \text{almost surely over } (G, \rho). \quad (1.1)$$

On the other hand, Kesten's approach only works for the *extrinsic* Euclidean metric, and not for the *intrinsic* metric  $d_G$  (which can be arbitrarily larger).

Kesten asked whether (1.1) holds for *any* (deterministic) subgraph  $G$  of  $\mathbb{Z}^d$ . Barlow and Perkins [BP89] answered this negatively: They exhibit a subgraph of  $\mathbb{Z}^2$  on which the Varopoulos-Carne bound is asymptotically tight (even for the Euclidean metric).

**Random walk and the growth of harmonic functions.** One motivation for studying situations in which Varopoulos-Carne can be improved comes from the theory of harmonic functions and their role in geometric analysis and in recent proofs of the central limit theorem for random graphs. Indeed, this led the authors of [BDCKY15] to study harmonic functions in random environments.

Consider a random rooted graph  $(G, \rho)$ . We will assume that  $G$  is locally finite and almost surely connected. Let  $\{X_n\}$  denote the random walk conditioned on  $(G, \rho)$ . Unless otherwise stated, we take  $X_0 = \rho$ .

**Definition 1.1.**  $(G, \rho)$  is said to be *stationary* if  $(G, X_0) \stackrel{\text{law}}{=} (G, X_1)$ .

Let  $V_G$  be the vertex set of  $G$ , and let  $d_G$  denote the graph metric on  $G$ . For  $x \in V_G$ , we use the notation

$$B_x^G(r) = \{y \in V_G : d_G(x, y) \leq r\}.$$

The random graph  $(G, \rho)$  has *annealed polynomial growth* if there exist constants  $c, d > 0$  such that for  $r \geq 1$ ,

$$\mathbb{E} [|B_\rho^G(r)|] \leq cr^d. \quad (1.2)$$

Say that the random walk  $\{X_n\}$  is *at most diffusive* if there is a constant  $C > 0$  such that

$$\mathbb{E} [d_G(X_0, X_n)^2] \leq Cn \quad (1.3)$$

for all  $n \geq 1$ .

We now state the main result of [BDCKY15] for the special case of stationary random graphs. A harmonic function conditioned on  $G$  is a map  $h : V_G \rightarrow \mathbb{R}$  satisfying

$$\mathbb{E}[h(X_1) \mid X_0 = x] = h(x) \quad \forall x \in V_G.$$

Say that  $h$  has *sublinear growth* if for every infinite sequence  $\{x_n\} \subseteq V_G$  with  $d_G(\rho, x_n) \rightarrow \infty$ , it holds that

$$\lim_{n \rightarrow \infty} \frac{|h(x_n)|}{d_G(\rho, x_n)} = 0.$$

**Theorem 1.2** ([BDCKY15]). *Suppose  $(G, \rho)$  is a stationary random graph with annealed polynomial growth, and suppose the random walk on  $(G, \rho)$  is at most diffusive in the sense of (1.3). Then almost surely  $G$  does not admit a non-constant harmonic function of sublinear growth.*

Our main result is that the diffusivity assumption can be removed. Say that  $(G, \rho)$  has *weakly annealed polynomial growth* if there are non-negative constants  $c, s \geq 0$  such that for  $r \geq 1$ ,

$$\mathbb{E} \left[ \log |B_\rho^G(r)| \right] \leq s \log r + c. \quad (1.4)$$

(Note that this is a *weaker* assumption than annealed polynomial growth.)

**Theorem 1.3.** *Suppose  $(G, \rho)$  is a stationary random graph with weakly annealed polynomial growth. Then almost surely  $G$  does not admit a non-constant harmonic function of sublinear growth.*

Our proof of [Theorem 1.3](#) proceeds in the natural way: We show that weakly annealed polynomial growth always yields a sequence of times at which the random walk is at most diffusive.

**Theorem 1.4.** *If  $(G, \rho)$  is a stationary random graph of annealed polynomial growth, then for every  $\varepsilon > 0$ , there is a constant  $C > 0$  and an infinite (deterministic) sequence of times  $\{t_n\}$  such that*

$$\mathbb{P} \left( \mathbb{E} \left[ d_G(X_0, X_{t_n})^2 \mid (G, \rho) \right] \leq C t_n \right) \geq 1 - \varepsilon.$$

Note that [Theorem 1.4](#) is new even for stationary random subgraphs of  $\mathbb{Z}^d$  since, in contrast to Kesten's work, we are able to bound the speed of the random walk in the intrinsic metric. To complement this result, we show that passing to a subsequence of times is necessary: There are stationary random graphs of (almost sure) polynomial growth on which the random walk is almost surely superdiffusive at an infinite subset of times.

**Theorem 1.5** (See [Theorem 4.1](#)). *There is a stationary random graph  $(G, \rho)$  of almost sure polynomial growth such that for an infinite (deterministic) sequence of times  $\{t_n\}$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \mathbb{E} \left[ d_G(X_0, X_{t_n})^2 \mid (G, \rho) \right] \geq t_n (\log t_n)^{0.9} \right) = 1.$$

We remark that instead of  $(\log t_n)^{0.9}$ , one could put  $f(t_n)$  for any function satisfying  $f(t) \leq o(\log t)$  as  $t \rightarrow \infty$ . This is almost tight as it nearly matches the Varopoulos-Carne estimate (see, e.g., [\[Woe00, Ch. 14\]](#)). Our work leaves open the intriguing question of whether [Theorem 1.4](#) holds for all times when  $(G, \rho)$  is a stationary random subgraph of  $\mathbb{Z}^d$ .

## 1.1 The absence of non-constant sublinear growth harmonic functions

Let us recall that the *entropy of  $X_n$  conditioned on  $(G, \rho)$* :

$$H(X_n \mid (G, \rho)) = \sum_{x \in V_G} \mathbb{P}[X_n = x \mid (G, \rho)] \log \frac{1}{\mathbb{P}[X_n = x \mid (G, \rho)]}, \quad (1.5)$$

with the convention that  $0 \log 0 = 0$ . Similarly we define  $H((X_1, X_n) \mid (G, \rho))$  to be the entropy of the joint distribution of  $(X_1, X_n)$ , conditioned on  $(G, \rho)$ . To simplify notation, we will denote  $H(X_n \mid (G, \rho))$  and  $H((X_1, X_n) \mid (G, \rho))$  by  $H_{(G, \rho)}(n)$  and  $H_{(G, \rho)}(1, n)$ , respectively.

Define the *annealed entropy* by

$$H_n = \mathbb{E}[H_{(G, \rho)}(n)]. \quad (1.6)$$

Our proof of [Theorem 1.3](#) is based on the main result of [\[BDCKY15\]](#) which exploits connections between harmonic functions and the escape rate of random walk on graphs. This reduces proving [Theorem 1.3](#) to proving the following.

**Theorem 1.6.** *If  $(G, \rho)$  is a stationary random graph of weakly annealed polynomial growth, then for every  $\varepsilon > 0$ , there is a constant  $C > 0$  and an infinite (deterministic) sequence of times  $\{t_n\}$  such that  $H_{t_n} - H_{t_{n-1}} \leq \frac{C}{t_n}$  and,*

$$\mathbb{P} \left( \mathbb{E} \left[ d_G(X_0, X_{t_n})^2 \mid (G, \rho) \right] \leq C t_n \right) \geq 1 - \varepsilon.$$

The proof of the preceding theorem constitutes the bulk of this article. We first show how [Theorem 1.3](#) follows.

*Proof of [Theorem 1.3](#).* Observe that by the chain rule for entropy and stationarity of  $(G, \rho)$ , it follows that for any  $t \geq 1$ ,

$$\mathbb{E}[H_{(G,\rho)}(1, t) - H_{(G,\rho)}(1)] = H_{t-1}. \quad (1.7)$$

There by [Theorem 1.6](#), (1.7) and Fatou's Lemma, for any  $\varepsilon > 0$ , there is a constant  $C > 0$  such that with probability  $1 - \varepsilon$  over the choice of  $(G, \rho)$ , there exists an infinite sequence of times  $\{t_n\}$  (depending on  $(G, \rho)$ ) such that,

$$H_{(G,\rho)}(t_n) + H_{(G,\rho)}(1) - H_{(G,\rho)}(1, t_n) \leq \frac{C}{t_n} \text{ and } \mathbb{E}[d_G(X_0, X_{t_n})^2 \mid (G, \rho)] \leq Ct_n, \quad (1.8)$$

(for the first inequality notice  $\mathbb{E}[H_{(G,\rho)}(t_n) + H_{(G,\rho)}(1) - H_{(G,\rho)}(1, t_n)] = H_{t_n} - H_{t_{n-1}}$ ).

Suppose  $h : V_G \rightarrow \mathbb{R}$  is harmonic on  $G$ . The authors of [\[BDCKY15\]](#) establish the inequality: For any time  $t \geq 1$ ,

$$\mathbb{E}[|h(X_0) - h(X_1)| \mid (G, \rho)] \leq \sqrt{4[H_{(G,\rho)}(1) + H_{(G,\rho)}(t) - H_{(G,\rho)}(1, t)] \cdot \mathbb{E}[h(X_t)^2 \mid (G, \rho)]}. \quad (1.9)$$

(This inequality is the conjunction of inequality (11) and the first inequality in the proof of Theorem 8 in [\[BDCKY15\]](#).)

If the graph  $(G, \rho)$  is such that (1.8) holds and  $h : V_G \rightarrow \mathbb{R}$  has sublinear growth, then if we consider (1.9) along the sequence  $t = t_n$  and send  $n \rightarrow \infty$ , we conclude that almost surely

$$h(X_0) = h(X_1).$$

Now send  $\varepsilon \rightarrow 0$  to conclude that almost surely on  $(G, \rho)$  and the random walk  $\{X_t\}$ , we have  $h(X_0) = h(X_1)$ .

By stationarity, this implies  $h(X_t) = h(X_{t+1})$  almost surely for every time  $t \geq 1$ . Since  $G$  is almost surely connected, we conclude that  $h$  must be constant. Therefore [Theorem 1.6](#) implies that almost surely  $(G, \rho)$  does not admit a non-constant harmonic function of sublinear growth.  $\square$

## 1.2 Speed via Euclidean embeddings

Note that for discrete groups of polynomial growth, significantly stronger results than [Theorem 1.4](#) are known (giving precise estimates on the heat kernel). See, for instance, the work of Hebisch and Saloff-Coste [\[HSC93\]](#). But those estimates require detailed information about the geometry that is furnished by Gromov's classification of such groups (in particular, they require the counting measure to be doubling). Clearly such methods are unavailable in our setting.

Even when one does not know that the counting measure is doubling, polynomial growth of a graph  $G$  still yields infinitely many radii  $r > 0$  at which  $|B_x^G(2r)| \leq C|B_x^G(r)|$  for some constant  $C > 0$  depending only on the growth rate. Indeed, locating such scales and performing geometric arguments that depend only on the local doubling constant underlie Kleiner's remarkable proof of Gromov's theorem [\[Kle10\]](#) (see also the quantitative results in [\[ST10\]](#)). (Somewhat related to the topic of the current paper, the heart of Kleiner's argument lies in establishing that on any finitely generated group of polynomial growth, the space of harmonic functions of (fixed) polynomial growth is finite-dimensional.)

We will pursue a related course, but in order to bound the speed of the random walk after  $n$  steps, we require control on the volume growth over  $\approx \log n$  scales, corresponding to distances in the interval  $[\sqrt{n}, n]$ . Polynomial volume growth is certainly not sufficient to find  $\log n$  consecutive scales at which the growth is doubling (uniformly in  $n$ ). Confronting this difficulty is the major technical challenge we face.

**Reducing to analysis on finite subgraphs.** In order to establish [Theorem 1.6](#), we first invoke the mass transport principle to show that it suffices to examine the random walk restricted to finite subgraphs of  $(G, \rho)$ . Let  $\mu_G(S) = \sum_{x \in S} \deg_G(x)$  for all subsets  $S \subseteq V_G$ .

In [Section 3.1](#), we argue that it is enough to find an infinite sequence of times  $\mathbb{T}$  and radii  $\{r_n : n \in \mathbb{T}\}$  such that the following three conditions hold for some constant  $C$ :

1. For every  $\varepsilon > 0$  and all  $n \in \mathbb{T}$  with  $n \geq (1/\varepsilon)^4$ ,

$$\mathbb{E} \left[ \frac{\mu_G \left( \left\{ x \in B_\rho^G(r_n) : \mathbb{E}[d_G(X_0, X_n)^2 \mid X_0 = x, (G, \rho)] \geq (C\varepsilon^{-13})n \right\} \right)}{\mu_G(B_\rho^G(r_n))} \right] \leq \varepsilon, \quad (1.10)$$

2.  $H_n - H_{n-1} \leq \frac{C}{n}$ ,

3.  $\mathbb{E} \left[ \log \frac{\mu_G(B_\rho^G(2r_n))}{\mu_G(B_\rho^G(r_n))} \right] \leq C$ .

It is noteworthy that our application of the mass transport principle uses the polynomial growth condition; specifically, we need to apply it at a scale where  $\mu_G$  is doubling (see [Lemma 3.1](#)).

**Embeddings and martingales.** Let us focus now on condition (1) since it is the difficult one to verify. In order to control the speed of the random walk started at a uniformly random point of  $B_\rho^G(r_n)$ , we construct a family of mappings  $\{F_k\}$  from  $B_\rho^G(r_n)$  into a Hilbert space and use the martingale methods of [\[NPSS06, DLP13\]](#) to derive bounds on the speed. The following statement is a slightly weaker version of [Lemma 2.3](#) in [Section 2.1](#).

**Lemma 1.7.** *Consider a graph  $G = (V_G, E_G)$ , a finite subset  $S \subseteq V_G$ , and a family  $\{F_k : S \rightarrow \mathcal{H}\}_{k \in \mathbb{N}}$  of 1-Lipschitz mappings into a Hilbert space. Let  $\varphi : \mathbb{N} \rightarrow \mathbb{R}_+$  be a given function. For  $k_0 \in \mathbb{N}$ , define the set of pairs*

$$\mathcal{G}(k_0, \varphi) = \left\{ (x, y) \in S^2 : \text{for all } k \geq k_0, d_G(x, y) \geq 8^k \implies \|F_k(x) - F_k(y)\|_{\mathcal{H}} \geq \frac{8^k}{\varphi(k)} \right\}.$$

If  $\{Z_t\}$  is the stationary random walk restricted to  $S$  (cf. [Definition 1.12](#)), then for every  $n \geq 1$ ,

$$\mathbb{E} \left[ d_G(Z_{2n}, Z_0)^2 \cdot \mathbf{1}_{\mathcal{G}(\alpha_n, \varphi)}(Z_0, Z_{2n}) \right] \leq 2n + 256 \sum_{k \geq \alpha_n} 8^{2k} \exp \left( \frac{-8^{2k}}{32n\varphi(k)^2} \right), \quad (1.11)$$

where  $\alpha_n = \lceil \log_8(\sqrt{2n}) \rceil$ .

In [Section 2.2](#), we show how standard tools from metric embedding theory [\[CKR01, KLMN05\]](#) provide a family of maps which are co-Lipschitz at a fixed scale, assuming the growth rate of balls at that scale is small.

**Lemma 1.8** (Statement of [Lemma 2.5](#)). *For any graph  $G = (V_G, E_G)$  and any  $k \geq 1$ , there is a 1-Lipschitz map  $F_k : V_G \rightarrow \ell_2$  such that for all  $x, y \in V_G$ , it holds that*

$$d_G(x, y) \geq 8^k \implies \|F_k(x) - F_k(y)\|_2 \geq \frac{8^k}{128 \left( 1 + \log \frac{|B_x^G(8^k)|}{|B_x^G(8^{k-1})|} \right)}$$

It may help to consider now the following special case: Suppose that the counting measure on  $G$  is doubling, i.e.

$$\sup_{x \in V_G, r \geq 0} \frac{|B_x^G(2r)|}{|B_x^G(r)|} \leq O(1).$$

In that case, if we use the family  $\{F_k\}$  from [Lemma 1.8](#), then there is some uniformly bounded function  $\varphi : \mathbb{N} \rightarrow \mathbb{R}_+$  in [Lemma 1.7](#) such that  $\mathbf{1}_{\mathcal{G}(k_0, \varphi)} \equiv \mathbf{1}$  for all  $k_0 \geq 1$ . Evaluating the sum in (1.11) immediately yields  $\mathbb{E}[d_G(Z_{2n}, Z_0)^2] \leq O(n)$ , completing our verification of (1.10). (Strictly speaking, the stationary measure on  $S$  and the measure  $\mu_G$  restricted to  $S$  are different, but they can be made arbitrarily close by taking  $S = B_\rho^G(r_n)$  where  $r_n$  is chosen so that  $S$  is a sufficiently good Følner set.)

In general, polynomial growth does not imply that the counting measure is doubling (and certainly the annealed form introduces even more complexity). Still, using [Lemma 1.7](#) and [Lemma 1.8](#) in conjunction, in [Section 3.2](#) we show that (1.10) holds at time  $n$  (for some radius  $r_n \gg n$ ) if the average profile of growth rates of balls  $\{B_x^G(r) \subseteq B_\rho^G(r_n)\}$  is sufficiently well-behaved for  $r \in [\sqrt{n}, n]$ .

Finally, in [Section 3.3](#), we argue that the annealed growth condition (1.4) allows us to find an infinite sequence of radii at which the average growth profile is well-behaved (with high probability over the choice of  $(G, \rho)$ ). This is subtle, as we require control on the growth for  $\approx \log n$  scales (corresponding to  $r \in [\sqrt{n}, n]$ ).<sup>1</sup> As mentioned before, one cannot hope to find such a sequence of consecutive scales at which the volume growth is uniformly doubling. Fortunately, the subgaussian tail in (1.11) gives us some flexibility; it will suffice to find a sequence of consecutive scales where the volume growth is not increasing too fast. Once this is established, we can verify (1.10) along this sequence and confirm [Theorem 1.6](#).

### 1.3 A deterministic example: Planar graphs

In this section, we present a solution to a question of Benjamini about random walks on planar graphs. It illustrates some of the ideas our main argument and their origins (in K. Ball's notion of Markov type), as well as the reduction of speed questions to the setting of stationary Markov chains on finite subgraphs.

Consider again a graph  $G = (V_G, E_G)$ . For a finite subset  $S \subseteq V_G$ , define the *edge boundary*

$$\partial_E S = \{ \{u, v\} \in E_G : \mathbf{1}_S(u) \neq \mathbf{1}_S(v) \},$$

and the *edge expansion* of  $S$  for  $S \neq \emptyset$ :

$$\phi_G(S) = \frac{|\partial_E S|}{\mu_G(S)}.$$

Say that  $G$  is *amenable* if  $\inf\{\phi_G(S) : \text{finite } S \subseteq V_G, S \neq \emptyset\} = 0$ . Otherwise, say that  $G$  is *non-amenable*.

Let  $\{X_t\}$  denote simple random walk on  $G$ . We say that the walk is *ballistic* if there is a constant  $c > 0$  such that for all  $v \in V_G$ ,

$$\mathbb{E} \left[ d_G(X_0, X_t)^2 \mid X_0 = v \right] \geq ct^2$$

for all  $t \geq 0$ . Say that the walk is *always somewhere at most diffusive* if there is a constant  $c > 0$  such that for all  $t \geq 0$ ,

$$\inf_{v \in V_G} \mathbb{E} \left[ d_G(X_0, X_t)^2 \mid X_0 = v \right] \leq ct.$$

The following result was conjectured by Itai Benjamini.<sup>2</sup> It states that for planar graphs, there are no intermediate (uniform) speeds between  $\sqrt{t}$  and  $t$ .

<sup>1</sup>The Varopoulos-Carne bound suggests we only need control for  $\log \log n$  scales corresponding to  $r \in [\sqrt{n}, \sqrt{n \log n}]$ , but the same problem arises.

<sup>2</sup>It was made by Benjamini at the Erdős Centennial in Budapest, July, 2013



**Theorem 1.9.** Suppose that  $G$  is an infinite planar graph with uniformly bounded vertex degrees. Either  $G$  is amenable and the random walk is always somewhere at most diffusive, or  $G$  is non-amenable and the random walk is ballistic.

Benjamini suggested this as an analog to the following dichotomy: Every amenable planar  $G$  graph admits arbitrarily large sets  $S \subseteq V_G$  such that  $|\partial_V S| \leq O(\sqrt{|S|})$ , where  $\partial_V S = \{v \in V_G : d_G(v, S) = 1\}$ . (This fact was announced by Gromov; see [Bow95] for a short proof.) Of course, in the non-amenable case, one has a linear isoperimetric profile:  $|\partial_V S| \geq c|S|$  for some  $c > 0$  and every  $S \subseteq V_G$ . Note that Theorem 1.9 is straightforward in the non-amenable case: If a graph  $G$  is non-amenable, then  $G$  has spectral radius  $\rho < 1$  [Kes59], hence the random walk is ballistic (see, e.g. [Woe00, Prop. 8.2]).

**Remark 1.10.** If one removes the assumption of bounded degrees from Theorem 1.9, then for  $G$  amenable, it still holds that the random walk is always somewhere at most diffusive (the argument below does not assume any bound on the vertex degrees). But there are non-amenable planar graphs for which the random walk does not have positive speed. We refer to [LP16, Ex 6.56] for a description of the unpublished construction of Angel, Hutchcroft, Nachmias, and Ray.

For the amenable case, we recall K. Ball's notion of Markov type [Bal92].

**Definition 1.11** (Markov type). A metric space  $(X, d)$  is said to have *Markov type*  $p \in [1, \infty)$  if there is a constant  $M > 0$  such that for every  $n \in \mathbb{N}$ , the following holds. For every reversible Markov chain  $\{Z_t\}_{t=0}^\infty$  on  $\{1, \dots, n\}$ , every mapping  $f : \{1, \dots, n\} \rightarrow X$ , and every time  $t \in \mathbb{N}$ ,

$$\mathbb{E} [d(f(Z_t), f(Z_0))^p] \leq M^p t \mathbb{E} [d(f(Z_0), f(Z_1))^p], \quad (1.12)$$

where  $Z_0$  is distributed according to the stationary measure of the chain. One denotes by  $M_p(X, d)$  the infimal constant  $M$  such that the inequality holds.

**Definition 1.12** (Restricted random walk). Consider a graph  $G = (V_G, E_G)$ , and let

$$N(x) = \{y \in V_G : \{x, y\} \in E_G\}$$

denote the neighborhood of a vertex  $x \in V_G$ . Fix a finite subset  $S \subseteq V_G$ . Denote the measure  $\pi$  on  $S$  by  $\pi(x) = \deg_G(x)/\mu_G(S)$ . We define the *random walk restricted to  $S$*  as the following process  $\{Z_t\}$ : For  $t \geq 0$ , put

$$\mathbb{P}(Z_{t+1} = y \mid Z_t = x) = \begin{cases} \frac{|N(x) \setminus S|}{\deg_G(x)} & y = x \\ \frac{1}{\deg_G(x)} & y \in N(x) \cap S \\ 0 & \text{otherwise.} \end{cases}$$

It is straightforward to check that  $\{Z_t\}$  is a reversible Markov chain on  $S$  with stationary measure  $\pi$ . If  $Z_0$  has law  $\pi$ , we say that  $\{Z_t\}$  is the *stationary random walk restricted to  $S$* .

**Definition 1.13** (Graphic Markov type). Define the *graphic Markov type  $p$  constant*  $M_p^{\text{gr}}(G)$  of a graph  $G = (V_G, E_G)$  as the infimal number  $M$  such that for every finite subset  $S \subseteq V_G$  and  $t \in \mathbb{N}$ ,

$$\mathbb{E} [d_G(Z_0^S, Z_t^S)^p] \leq M^p t,$$

where  $\{Z_t^S\}$  is the stationary random walk restricted to  $S$ .

**Lemma 1.14.** If  $G$  is an amenable graph, then for every time  $t \geq 0$ ,

$$\inf_{v \in V_G} \mathbb{E} [d_G(X_0, X_t)^2 \mid X_0 = v] \leq 2 \left( M_2^{\text{gr}}(G) \right)^2 t \leq 2 M_2(G)^2 t.$$

We will prove this lemma momentarily. Let us observe first that [Theorem 1.9](#) follows immediately in conjunction with the next theorem.

**Theorem 1.15** ([\[DLP13\]](#)). *There is a constant  $K > 0$  such that  $M_2(G) \leq K$  for any planar graph  $G$ .*

We remark that bounding  $M_2^{\text{gr}}(G)$  (which is all that is needed to apply [Lemma 1.14](#)) is somewhat easier than bounding  $M_2(G)$ ; see [Corollary 2.4](#) and the remarks thereafter.

*Proof of [Lemma 1.14](#).* Fix a subset  $S \subseteq V_G$ . Let  $\{Z_t\}$  denote the stationary random walk restricted to  $S$ . From the definition of graphic Markov type, for every  $t \geq 0$ , we have

$$\mathbb{E}[d_G(Z_0, Z_t)^2] \leq (M_2^{\text{gr}}(G))^2 t \mathbb{E}[d_G(Z_0, Z_1)^2] \leq (M_2^{\text{gr}}(G))^2 t. \quad (1.13)$$

Note that since  $Z_t$  is stationary, it holds that for all  $t \geq 0$ , we have  $\mathbb{P}(Z_{t+1} = Z_t) = \phi_G(S)$ . Recall that  $\{X_t\}$  is the random walk on  $G$ . If  $X_0$  has the law of  $Z_0$ , then  $X_t$  has the law of  $Z_t$  conditioned on the event  $\{X_0, X_1, \dots, X_t\} \subseteq S$ .

In particular, we can conclude that

$$\mathbb{P}[\{X_0, X_1, \dots, X_t\} \not\subseteq S \mid X_0 = Z_0] \leq \sum_{i=0}^{t-1} \mathbb{P}(Z_{i+1} = Z_i) \leq t\phi_G(S). \quad (1.14)$$

Hence,

$$\begin{aligned} \mathbb{E}[d_G(X_t, X_0)^2 \mid X_0 = Z_0] &\leq \mathbb{E}[d_G(Z_t, Z_0)^2] + \mathbb{P}[\{X_0, X_1, \dots, X_t\} \not\subseteq S \mid X_0 = Z_0] \cdot t^2 \\ &\leq (M_2^{\text{gr}}(G))^2 t + t^3 \phi_G(S), \end{aligned} \quad (1.15)$$

where in the first line we have used the fact that  $d_G(X_t, X_0)^2 \leq t^2$  holds with probability one, and in the second line we have employed the bounds [\(1.13\)](#) and [\(1.14\)](#).

Now fix a time  $t \geq 0$ . Since  $G$  is amenable, there exists a choice of  $S$  for which  $\phi_G(S) \leq (M_2^{\text{gr}}(G)/t)^2$ . In this case, from [\(1.15\)](#) we obtain

$$\mathbb{E}[d_G(X_t, X_0)^2 \mid X_0 = Z_0] \leq 2 (M_2^{\text{gr}}(G))^2 t.$$

Thus certainly the bound holds for some fixed  $X_0 \in S$ , concluding the proof.  $\square$

## 2 Martingales, embeddings, and growth rates

Our proof of [Theorem 1.6](#) involves the construction of embeddings of  $(G, \rho)$  into a Hilbert space  $\mathcal{H}$ . The embeddings give rise to a family of martingales in  $\mathcal{H}$  whose behavior can be used to control the speed of the random walk in  $G$ . This section is primarily expository; we review the martingale methods of [\[NPSS06, DLP13\]](#) and a construction of Euclidean embeddings that reflect the local geometry of a discrete metric space at a fixed scale [\[CKR01, KLMN05\]](#).

### 2.1 Control by martingales

Consider a finite metric space  $(X, d)$ . Let  $\{Z_t\}$  denote a stationary, reversible Markov chain on  $X$  with the property that

$$d(Z_0, Z_1) \leq 1 \quad \text{almost surely.} \quad (2.1)$$

Let  $\mathcal{Y}$  be a normed space and for a map  $f : X \rightarrow \mathcal{Y}$ , define

$$\|f\|_{\text{Lip}} = \max_{x \neq y \in X} \frac{\|f(x) - f(y)\|_{\mathcal{Y}}}{d(x, y)}.$$



The following result is proved in [NPSS06] (see also [LZ94]). A similar decomposition appears already in the work of Kesten [Kes86] (see the discussion in [BP89, Sec. 2]) for the special case of percolation clusters in  $\mathbb{Z}^d$ . A stark difference is that in Kesten's paper, the Markov chain  $\{Z_t\}$  already takes values in a subset of  $\mathbb{Z}^d$  (and hence the map  $f$  does not appear). On the other hand, this means that Kesten only bounds the speed of the walk in the ambient Euclidean metric, whereas we are interested in the speed in the intrinsic metric (which is larger, and hence harder to bound from above).

**Lemma 2.1.** *Then for every  $n \geq 1$ , there is a forward martingale  $\{A_t\}$  and a backward martingale  $\{B_t\}$  such that*

1.  $f(Z_{2n}) - f(Z_0) = A_n - B_n$
2. For all  $t = 1, 2, \dots, n$ , it holds that

$$\|A_t - A_{t-1}\|_{\mathcal{Y}}, \|B_t - B_{t-1}\|_{\mathcal{Y}} \leq 2\|f\|_{\text{Lip}}.$$

For completeness we include the proof.

*Proof.* Define the martingales  $\{M_s\}_{s=0}^{2n}$  and  $\{N_s\}_{s=0}^{2n}$  by  $M_0 = f(Z_0)$  and  $N_0 = f(Z_{2n})$  and for  $0 \leq s \leq t-1$ ,

$$\begin{aligned} M_{s+1} - M_s &:= f(Z_{s+1}) - f(Z_s) - \mathbb{E}[f(Z_{s+1}) - f(Z_s) \mid Z_s] \\ N_{s+1} - N_s &:= f(Z_{t-s-1}) - f(Z_{t-s}) - \mathbb{E}[f(Z_{t-s-1}) - f(Z_{t-s}) \mid Z_{t-s}]. \end{aligned} \quad (2.2)$$

Observe that  $\{M_s\}$  is a martingale with respect to the filtration induced on  $\{Z_0, Z_1, \dots, Z_{2n}\}$  and  $\{N_s\}$  is a martingale with respect to the filtration induced on  $\{Z_{2n}, Z_{2n-1}, \dots, Z_0\}$ .

For every  $1 \leq s \leq 2n-1$  using stationarity gives  $\mathbb{E}[f(Z_{s+1}) \mid Z_s] = \mathbb{E}[f(Z_{s-1}) \mid Z_s]$ , and it follows that

$$f(Z_{s+1}) - f(Z_{s-1}) = (M_{s+1} - M_s) - (N_{2n-s+1} - N_{2n-s}). \quad (2.3)$$

Now consider the martingales  $\{A_t\}_{0 \leq t \leq n}$  and  $\{B_t\}_{0 \leq t \leq n}$  given by

$$\begin{aligned} A_t &:= \sum_{s=0}^t M_{2s} - M_{2s-1} \\ B_t &:= \sum_{s=0}^t N_{2s} - N_{2s-1}. \end{aligned}$$

(2) follows the preceding definition and (2.2), along with assumption (2.1). The proof of (1) is by summing (2.3) over  $s = 1, 3, \dots, 2n-1$ .  $\square$

Combining Lemma 2.1 with Azuma's inequality for  $\mathcal{H}$ -valued martingales [Pin94] yields the following.

**Corollary 2.2.** *If  $\mathcal{H}$  is a Hilbert space, then for all  $n \geq 1$ ,*

$$\mathbb{P}(\|f(Z_{2n}) - f(Z_0)\|_{\mathcal{H}} \geq \lambda) \leq 4 \exp\left(\frac{-\lambda^2}{32n\|f\|_{\text{Lip}}^2}\right)$$

Define the constants

$$\alpha_n = \lceil \log_8(\sqrt{2n}) \rceil \quad (2.4)$$

$$\beta_n = \lceil \log_8(2n) \rceil. \quad (2.5)$$

**Lemma 2.3.** Consider a graph  $G = (V_G, E_G)$ , a finite subset  $S \subseteq V_G$ , and a family  $\{F_k : S \rightarrow \mathcal{H}\}_{k \in \mathbb{N}}$  of 1-Lipschitz mappings into a Hilbert space. Let  $\varphi : \mathbb{N} \rightarrow \mathbb{R}_+$  be a given function. For  $k_0 \in \mathbb{N}$ , define the set of pairs

$$\mathcal{G}(k_0, \varphi) = \left\{ (x, y) \in S^2 : \text{for all } k \geq k_0, d_G(x, y) \geq 8^k \implies \|F_k(x) - F_k(y)\|_{\mathcal{H}} \geq \frac{8^k}{\varphi(k)} \right\}.$$

If  $\{Z_t\}$  is the stationary random walk restricted to  $S$  (cf. [Definition 1.12](#)), then for every  $n \geq 1$ ,

$$\mathbb{E} \left[ d_G(Z_{2n}, Z_0)^2 \cdot \mathbf{1}_{\mathcal{G}(\alpha_n, \varphi)}(Z_0, Z_{2n}) \right] \leq 2n + 256 \sum_{k=\alpha_n}^{\beta_n} 8^{2k} \exp \left( \frac{-8^{2k}}{32n\varphi(k)^2} \right).$$

*Proof.* Use that fact that for a non-negative random variable  $X$ , we have  $\mathbb{E}[X^2] \leq \sum_{k=0}^{\infty} 8^{2k+2} \mathbb{P}[X \geq 8^k]$  to write

$$\begin{aligned} \mathbb{E}[d_G(Z_{2n}, Z_0)^2 \cdot \mathbf{1}_{\mathcal{G}(\alpha_n, \varphi)}(Z_0, Z_{2n})] &\leq \sum_{k=0}^{\beta_n} 8^{2k+2} \mathbb{P}[d_G(Z_{2n}, Z_0) \geq 8^k \wedge (Z_0, Z_{2n}) \in \mathcal{G}(\alpha_n, \varphi)] \\ &\leq \sum_{k=0}^{\beta_n} 8^{2k+2} \mathbb{P} \left[ \|F_k(Z_0) - F_k(Z_{2n})\|_{\mathcal{H}} \geq \frac{8^k}{\varphi(k)} \right] \\ &\leq 2n + \sum_{k=\alpha_n}^{\beta_n} 8^{2k+2} \mathbb{P} \left[ \|F_k(Z_0) - F_k(Z_{2n})\|_{\mathcal{H}} \geq \frac{8^k}{\varphi(k)} \right], \end{aligned}$$

where in the first inequality we have used the fact that  $d_G(Z_0, Z_{2n}) \leq 2n$  is always true. The desired bound now follows from [Corollary 2.2](#).  $\square$

We remark on one straightforward (but illustrative) application of [Lemma 2.3](#). Following [\[DLP13\]](#), we say that a metric space  $(X, d)$  admits a *threshold embedding with distortion  $D$  into a Hilbert space  $\mathcal{H}$*  if there is a family of 1-Lipschitz maps  $\{F_k : X \rightarrow \mathcal{H}\}$  such that

$$x, y \in X \text{ and } d(x, y) \geq 8^k \implies \|F_k(x) - F_k(y)\|_{\mathcal{H}} \geq \frac{8^k}{D}. \quad (2.6)$$

It is proved in [\[DLP13\]](#) that if such a threshold embedding exists, then  $M_2(X, d) \leq O(D)$  (recall the definition of Markov type from [Section 1.3](#)). Bounding the graphic Markov type is substantially easier.

**Corollary 2.4.** If  $G = (V_G, E_G)$  is a graph and  $(V_G, d_G)$  admits a threshold embedding into a Hilbert space  $\mathcal{H}$  with distortion  $D$ , then

$$M_2^{\text{gr}}(G) \leq O(D).$$

*Proof.* Fix a finite subset  $S \subseteq V_G$ . Let  $\{Z_t\}$  denote the stationary random walk restricted to  $S$ . Let  $\{F_k : V_G \rightarrow \mathcal{H}\}$  be the claimed threshold embedding. Apply [Lemma 2.3](#) to the family  $\{F_k|_S\}$  with  $\varphi \equiv D$ , in which case  $\mathbf{1}_{\mathcal{G}(\alpha_n, k_0)} \equiv \mathbf{1}$ . One concludes that for every  $n \geq 1$ ,

$$\mathbb{E}[d_G(Z_{2n}, Z_0)^2] \leq O(Dn).$$

Using  $d_G(Z_n, Z_0) \leq d_G(Z_{n+1}, Z_0) + 1$  yields a similar estimate for odd times, completing the proof.  $\square$

On the other hand, we will not have a uniform lower bound as in (2.6) that holds for all pairs  $x, y \in X$ .

**Volume growth.** Let  $G$  be a graph with vertex set  $V_G$ . For  $x \in V_G$ , we recall that  $B_x^G(R)$  is the closed  $R$ -ball around  $x$  in the metric  $d_G$ . Define

$$\phi_x^G(k) = \log \frac{|B_x^G(8^k)|}{|B_x^G(8^{k-1})|}. \quad (2.7)$$

In the next section, we exhibit a family of mappings that reflect the geometry of  $G$  well at scale  $8^k$  when  $\phi_x^G(k)$  is small.

**Lemma 2.5.** *For any  $k \geq 1$ , there is a 1-Lipschitz map  $F_k : V_G \rightarrow \ell_2$  such that for all  $x, y \in V_G$ , it holds that*

$$d_G(x, y) \geq 8^k \implies \|F_k(x) - F_k(y)\|_2 \geq \frac{8^k}{128(1 + \phi_x^G(k))}$$

## 2.2 Embeddings and growth rates

For a metric space  $(X, d)$ , define  $B(x, R) = \{y \in X : d(x, y) \leq R\}$ . We now prove the following generalization of [Lemma 2.5](#).

**Lemma 2.6.** *If  $(X, d)$  is a discrete metric space, then the following holds. For any  $\tau > 0$ , there is a 1-Lipschitz mapping  $\varphi_\tau : X \rightarrow \ell_2$  such that for all  $x, y \in X$ ,*

$$d(x, y) \geq \tau \implies \|\varphi_\tau(x) - \varphi_\tau(y)\|_2 \geq \frac{\tau}{128 \log \left( \frac{e|B(x, \frac{5}{8}\tau)|}{|B(x, \frac{1}{8}\tau)|} \right)}$$

[Lemma 2.5](#) is a well-known result in metric embedding theory; see, e.g., [\[KLMN05\]](#) where a similar lemma is stated. We provide a proof here for the sake of completeness.

By a simple compactness argument, it suffices to prove [Lemma 2.6](#) for  $X$  finite, which we now assume. Given a probability space  $(\Omega, \mathcal{B}, \mu)$ , we use  $L^2(\mu)$  to denote the Hilbert space of measurable real-valued random variables with inner product  $\langle Y, Z \rangle_{L^2(\mu)} = \mathbb{E}[YZ]$ . If  $P$  is a partition of  $X$ , we denote by  $P : X \rightarrow 2^X$  the map that sends  $x \in X$  to the unique set  $P(x) \in P$  containing  $x$ .

**Lemma 2.7.** *For any value  $\tau > 0$  and  $\varepsilon : X \rightarrow \mathbb{R}_+$ , the following holds. Let  $\mathcal{P}$  be a random partition of  $X$  with the following two properties:*

1. *Almost surely,  $\max_{S \in \mathcal{P}} \text{diam}_{(X, d)}(S) < \tau$ .*
2. *For every  $x \in X$ ,*

$$\mathbb{P}[B(x, \varepsilon(x)\tau) \subseteq \mathcal{P}(x)] \geq \delta.$$

*Then there exists a 1-Lipschitz mapping  $\varphi_\tau : X \rightarrow \ell_2$  such that for all  $x, y \in X$ ,*

$$d(x, y) \geq \tau \implies \|\varphi_\tau(x) - \varphi_\tau(y)\|_2 \geq \frac{\sqrt{\delta}}{2} \varepsilon(x)\tau.$$

*Proof.* For every  $P \in \text{supp}(\mathcal{P})$ , let  $\{\alpha_S : S \in P\}$  be a sequence of i.i.d. Bernoulli  $\{0, 1\}$  random variables (independent of  $\mathcal{P}$ ).

Consider the (random) map  $F : X \rightarrow \mathbb{R}$  given by

$$F(x) = \alpha_{P(x)} \cdot d(x, X \setminus P(x)).$$

By construction,  $F$  is almost surely 1-Lipschitz.

Now fix  $x, y \in X$  with  $d(x, y) \geq \tau$ . Note that by assumption (1),  $P(x) \neq P(y)$ . Therefore

$$\mathbb{E}|F(x) - F(y)|^2 \geq \varepsilon(x)^2 \tau^2 \cdot \mathbb{P}[B(x, \varepsilon(x)\tau) \subseteq P(x)] \cdot \mathbb{P}[\alpha_{P(x)} = 1] \cdot \mathbb{P}[\alpha_{P(y)} = 0] \geq \frac{\delta \varepsilon(x)^2 \tau^2}{4}.$$

Therefore  $F : X \rightarrow L^2(\mu)$  provides the desired mapping, where  $\mu$  is the law of the random map  $F$ . Note that since  $X$  is finite,  $\mu$  is finitely supported, so one can take  $L^2(\mu)$  as a finite-dimensional Hilbert space.  $\square$

In light of [Lemma 2.5](#), in order to prove [Lemma 2.6](#), it suffices to construct an appropriate random partition. To do so, we employ the method and analysis of [\[CKR01\]](#).

**Lemma 2.8.** *For every  $\tau > 0$ , there is a random partition  $\mathcal{P}$  satisfying the assumptions of [Lemma 2.5](#) with  $\delta = \frac{1}{2}$  and*

$$\varepsilon(x) = \left( 32 \log \left( \frac{e|B(x, \frac{5}{8}\tau)|}{|B(x, \frac{1}{8}\tau)|} \right) \right)^{-1}. \quad (2.8)$$

*Proof.* Suppose that  $|X| = n$  and let  $\pi : [n] \rightarrow X$  be a uniformly random bijection. Choose  $R \in [\frac{\tau}{4}, \frac{\tau}{2}]$  uniformly at random.

Let  $\mathcal{P}$  be the random partition constructed by iteratively cutting out the balls  $B(\pi(1), R), B(\pi(2), R), \dots, B(\pi(n), R)$ . In other words,  $\mathcal{P} = \{S_1, S_2, \dots, S_n\}$  where

$$S_i = B(\pi(i), R) \setminus (S_1 \cup \dots \cup S_{i-1}).$$

Fix a number  $\varepsilon \leq 1/8$  and a point  $x \in X$ . Let  $T \in [n]$  denote the smallest index for which  $d(\pi(T), x) \leq \varepsilon\tau + R$ . Then we have

$$\mathbb{P}[B(x, \varepsilon\tau) \not\subseteq \mathcal{P}(x)] \leq \mathbb{P}[d(\pi(T), x) \geq R - \varepsilon\tau]. \quad (2.9)$$

For  $y \in X$ , define the interval  $I(y) = [d(x, y) - \varepsilon\tau, d(x, y) + \varepsilon\tau]$ . Note that the bad event  $\{d(\pi(T), x) > R - \varepsilon\tau\}$  is the same as the event  $\{R \in I(\pi(T))\}$ .

Order the points of  $X$  in non-decreasing order from  $x$ :  $x_0 = x, x_1, \dots, x_n$ . Then [\(2.9\)](#) yields

$$\begin{aligned} \mathbb{P}[B(x, \varepsilon\tau) \not\subseteq \mathcal{P}(x)] &\leq \mathbb{P}[R \in I(\pi(T))] \\ &= \sum_{j=1}^n \mathbb{P}[R \in I(x_j)] \cdot \mathbb{P}[x_j = \pi(T) \mid R \in I(x_j)] \\ &\leq 16\varepsilon \sum_{j=|B(x, \tau/8)|+1}^{|B(x, \frac{5}{8}\tau)|} \mathbb{P}[x_j = \pi(T) \mid R \in I(x_j)] \end{aligned} \quad (2.10)$$

$$\begin{aligned} &\leq 16\varepsilon \sum_{j=|B(x, \tau/8)|+1}^{|B(x, \frac{5}{8}\tau)|} \frac{1}{j} \\ &\leq 16\varepsilon \log \left( \frac{e|B(x, \frac{5}{8}\tau)|}{|B(x, \frac{1}{8}\tau)|} \right). \end{aligned} \quad (2.11)$$

Inequality [\(2.10\)](#) arises from the fact that the length of  $I(x_j)$  is  $4\varepsilon\tau$  and  $R$  is chosen uniformly from an interval of length  $\tau/4$  and that if  $d(x, y) \leq \tau/8$  or  $d(x, y) > \frac{5}{8}\tau$ , then  $\mathbb{P}(R \in I(y)) = 0$ . Finally, to confirm [\(2.11\)](#), note that

$$R \in I(x_j) \implies R > d(x, x_j) - \varepsilon\tau \implies R > d(x, x_i) - \varepsilon\tau \text{ for } i \leq j.$$

In particular, conditioned on  $R \in I(x_j)$ , the event  $x_j = \pi(T)$  can only happen if  $x_j$  is chosen first from  $\{x_1, \dots, x_j\}$  in the permutation  $\pi$ .

Setting  $\varepsilon$  as in [\(2.8\)](#) completes the proof.  $\square$

### 3 Diffusive estimates

In order to apply the techniques of the preceding section, we need to reduce our main diffusive estimate (Theorem 1.6) to a statement about the random walk restricted to finite subgraphs in  $(G, \rho)$ . In Section 3.1, we use the mass transport principle to show that it suffices to control the speed of the random walk on an appropriate sequence of balls  $\{B_\rho^G(r_n)\}$  in  $G$ .

In Section 3.2, we argue that this is possible, conditioned on  $(G, \rho)$ , as long as there are good enough bounds on the average growth rate of balls  $\{B_x^G(r) \subseteq B_\rho^G(r_n)\}$ , where the average is taken over the stationary measure of the random walk restricted to  $B_\rho^G(r_n)$ . Finally, in Section 3.3, we show that the weakly annealed polynomial growth property shows yields an infinite sequence of radii  $\{r_n\}$  such that the average growth is controlled with high probability over the choice of  $(G, \rho)$ . This allows us to complete the proof of Theorem 1.6.

#### 3.1 The mass transport principle

We now return to the setting where  $(G, \rho)$  is a stationary random graph with vertex set  $V_G$ . For a subset  $S \subseteq V_G$ , define  $\mu_G(S) = \sum_{x \in S} \deg_G(x)$ .

In order to establish Theorem 1.6, we employ an unpublished result of Russ Lyons that every stationary random graph of (weakly) annealed subexponential growth is actually a reversible random graph. For completeness, we indicate a proof at the end of this section.

In particular, we can assume that  $(G, \rho)$  satisfies a mass transport principle (see, e.g., the extensive reference [AL07] or the discussion in [BC12]): For every positive functional  $F(G, \rho, x)$ , it holds that

$$\mathbb{E} \left[ \frac{1}{\deg_G(\rho)} \sum_{x \in V_G} F(G, \rho, x) \right] = \mathbb{E} \left[ \frac{1}{\deg_G(\rho)} \sum_{x \in V_G} F(G, x, \rho) \right]. \quad (3.1)$$

Consider an event  $\mathcal{A}$  in  $\mathcal{G}_\bullet$  (depending only on the isomorphism classes of finite rooted subgraphs).

**Lemma 3.1.** *For any  $R \geq 1$ , it holds that,*

$$\mathbb{E} \left[ \frac{\mu_G(B_\rho^G(R))}{\mu_G(B_\rho^G(2R))} \mathcal{A}(G, \rho) \right] \leq \mathbb{E} \left[ \frac{\mu_G(\{x \in B_\rho^G(R) : \mathbf{1}_{\mathcal{A}}(G, x)\})}{\mu_G(B_\rho^G(R))} \right].$$

*Proof.* Define a mass transportation:

$$F(G, \rho, x) = \deg_G(\rho) \frac{\deg_G(x)}{\mu_G(B_\rho^G(R))} \mathbf{1}_{B_\rho^G(R)}(x) \mathbf{1}_{\mathcal{A}}(G, x).$$

Observe that,

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{\mu_G(B_\rho^G(R))} \sum_{x \in B_\rho^G(R)} \deg_G(x) \mathbf{1}_{\mathcal{A}}(G, x) \right] &= \mathbb{E} \left[ \frac{1}{\deg_G(\rho)} \sum_{x \in V_G} F(G, \rho, x) \right] \\ &\stackrel{(3.1)}{=} \mathbb{E} \left[ \frac{1}{\deg_G(\rho)} \sum_{x \in V_G} F(G, x, \rho) \right] \\ &= \mathbb{E} \left[ \sum_{x \in B_\rho^G(R)} \frac{\deg_G(x)}{\mu_G(B_\rho^G(R))} \mathbf{1}_{\mathcal{A}}(G, \rho) \right] \end{aligned}$$

$$\geq \mathbb{E} \left[ \frac{\mu_G(B_\rho^G(R))}{\mu_G(B_\rho^G(2R))} \mathcal{A}(G, \rho) \right],$$

where the last line follows from the fact that  $x \in B_\rho^G(R) \implies \mu_G(B_x^G(R)) \leq \mu_G(B_\rho^G(2R))$ .  $\square$

The following theorem, along with the mass transport principle, implies [Theorem 1.6](#). Its proof occupies [Section 3.2](#) and [Section 3.3](#).

**Theorem 3.2.** *Suppose that  $(G, \rho)$  is a stationary random graph of weakly annealed polynomial growth. Then there is a constant  $C$  depending only on the growth constants of  $(G, \rho)$  (cf. (1.4)) and an infinite (deterministic) sequence of times  $\mathbb{T}$  and radii  $\{r_n : n \in \mathbb{T}\}$  such that the following conditions hold:*

1. *For every  $\varepsilon > 0$  and all  $n \in \mathbb{T}$  with  $n \geq (1/\varepsilon)^4$ , it holds that,*

$$\mathbb{E} \left[ \frac{\mu_G \left( \left\{ x \in B_\rho^G(r_n) : \mathbb{E}[d_G(X_0, X_n)^2 \mid X_0 = x, (G, \rho)] \geq (C\varepsilon^{-13})n \right\} \right)}{\mu_G(B_\rho^G(r_n))} \right] \leq \varepsilon,$$

2.  $(H_n - H_{n-1})n \leq C$ ,

3.  $\mathbb{E} \left[ \log \frac{\mu_G(B_\rho^G(2r_n))}{\mu_G(B_\rho^G(r_n))} \right] \leq C$ .

We finish off this section with the proof of [Theorem 1.6](#).

*Proof of Theorem 1.6.* Fix  $\delta > 0$  and apply [Theorem 3.2](#). Applying Markov's inequality to (3) yields

$$\mathbb{P} \left[ \log \frac{\mu_G(B_\rho^G(2r_n))}{\mu_G(B_\rho^G(r_n))} \leq \frac{C}{\delta} \right] \geq 1 - \delta.$$

We now lower bound the probability that the random walk started from the root is at most diffusive. Note that [Theorem 3.2](#) asserts this for the majority of the points in  $B_\rho^G(r_n)$ . To transfer this to the root, we use the mass transport principle.

To apply [Lemma 3.1](#), we define the set  $\mathcal{A}$  of rooted graphs such that

$$\mathbf{1}(\mathcal{A}) = \mathbf{1} \left( \left\{ (G, \rho) : \mathbb{E}_{X_n}[d_G(\rho, X_n)^2] \geq C'n \right\} \right)$$

for some constant  $C'$  which is specified below. Using [Lemma 3.1](#) in conjunction with (1) yields

$$e^{-C/\delta} \mathbb{P} \left( \left\{ \log \frac{\mu_G(B_\rho^G(2r_n))}{\mu_G(B_\rho^G(r_n))} \leq \frac{C}{\delta} \right\} \wedge \mathcal{A} \right) \leq \varepsilon.$$

Thus by union bound,

$$\mathbb{P}(\mathcal{A}) \leq e^{C/\delta} \varepsilon + \mathbb{P} \left( \log \frac{\mu_G(B_\rho^G(2r_n))}{\mu_G(B_\rho^G(r_n))} \geq \frac{C}{\delta} \right) \leq e^{C/\delta} \varepsilon + \delta.$$

Choosing  $\varepsilon = \delta e^{-C/\delta}$ , yields that for some  $C' \leq e^{14C/\delta}$ , and for all  $n \in \mathbb{T}$  sufficiently large,

$$\mathbb{P} \left( \mathbb{E}[d_G(X_0, X_n)^2 \mid (G, \rho)] \geq C'n \right) \leq \delta + e^{C/\delta} \delta e^{-C/\delta} \leq 2\delta.$$

Therefore it holds that for all  $n \in \mathbb{T}$  sufficiently large,  $H_n - H_{n-1} \leq \frac{C}{n}$  and,

$$\mathbb{P} \left[ \mathbb{E}[d_G(X_0, X_n)^2 \mid (G, \rho)] \geq C'n \right] \leq 3\delta,$$

yielding the desired result.  $\square$



### 3.1.1 Subexponential growth and reversibility

We now prove the following unpublished result of Russ Lyons.

Recall that  $(G, \rho)$  is stationary if  $(G, X_0) \stackrel{\text{law}}{=} (G, X_1)$  where  $\{X_n\}$  is the random walk on  $G$  with  $X_0 = \rho$ . The random graph  $(G, \rho)$  is said to be *reversible* if  $(G, X_0, X_1) \stackrel{\text{law}}{=} (G, X_1, X_0)$ .

**Theorem 3.3.** *If  $(G, \rho)$  is a stationary random graph such that*

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}[\log |B_\rho^G(n)|]}{n} \rightarrow 0, \quad (3.2)$$

*then  $(G, \rho)$  is reversible.*

This result was proved earlier in [BC12] with the additional assumption that  $\deg_G(\rho) \leq O(1)$  almost surely.

*Proof.* We will borrow heavily from [BC12, Ch. 4]. The reader is encouraged to consult that paper for more detailed explanations. Let  $\mu_{\rightarrow}$  and  $\mu_{\leftarrow}$  denote the laws of  $(G, X_0, X_1)$  and  $(G, X_1, X_0)$ , respectively. For a fixed graph  $G_0 = (V_0, E_0)$  and  $\{x, y\} \in E_0$ , we denote the Radon-Nikodym derivative

$$\Delta(G_0, x, y) := \frac{d\mu_{\leftarrow}}{d\mu_{\rightarrow}}(G_0, x, y).$$

One can extend this to pairs  $x, y \in V_0$  which are not necessarily adjacent: Consider any path  $x = x_0, x_1, \dots, x_n = y$  and define

$$\Delta(G_0, x, y) := \prod_{i=0}^{n-1} \Delta(G_0, x_i, x_{i+1}).$$

This value is independent of the path between  $x$  and  $y$  (see [BC12, Lem. 4.2]; this is a manifestation of the fact that a cycle and its reverse have the same probability under random walk on a graph).

Note that, because of this, for pairs  $x, y \in V_0$  such that  $\mathbb{P}[X_n = y \mid X_0 = x] > 0$ , it holds that

$$\Delta(G_0, x, y) = \frac{d\mu_{n,\leftarrow}}{d\mu_{n,\rightarrow}}(G_0, x, y), \quad (3.3)$$

where  $\mu_{n,\rightarrow}$  and  $\mu_{n,\leftarrow}$  are the laws of  $(G, X_0, X_n)$  and  $(G, X_n, X_0)$ , respectively. (This equality only makes sense up to sets of  $\mu_{\leftarrow}$ -measure zero.)

One has  $\mathbb{E}[\Delta(G, X_0, X_1)] = 1$ , and moreover Jensen's inequality shows that

$$\mathbb{E}[\log(\Delta(G, X_0, X_1))] \geq 0 \iff \Delta(G, X_0, X_1) = 1 \text{ a.s.} \iff (G, X_0, X_1) \text{ is reversible.} \quad (3.4)$$

Let  $\mathcal{G}_{\bullet\bullet}$  denote the set of isomorphism classes of bi-rooted graphs. Then for any Borel set  $A \subseteq \mathcal{G}_{\bullet\bullet}$  and  $n \geq 0$ , stationarity yields

$$\mathbb{P}[(G, X_0, X_n) \in A] = \mathbb{P}[(G, X_n, X_{2n}) \in A] \geq \mathbb{P}(X_{2n} = X_0 \mid X_n) \mathbb{P}[(G, X_n, X_0) \in A]. \quad (3.5)$$

Let  $p_G^n$  denote the  $n$ -step transition kernel in  $G$ . Using (3.5) in (3.3) implies that almost surely:

$$\Delta(G, X_0, X_n) \geq \mathbb{P}(X_{2n} = X_0 \mid X_n, G) = p_G^n(X_n, X_0) = p_G^n(X_0, X_n) \frac{\deg_G(X_0)}{\deg_G(X_n)}. \quad (3.6)$$

Therefore almost surely,

$$\mathbb{E}[\log \Delta(G, X_0, X_n) \mid G, X_0] \geq \sum_y p_G^n(X_0, y) \left( \log p_G^n(X_0, y) + \log \frac{\deg_G(X_0)}{\deg_G(y)} \right).$$

Note that (3.2) implies  $\mathbb{E}[\log \deg_G(\rho)] < \infty$ . Therefore taking expectations and again employing stationarity yields

$$\mathbb{E}[\log \Delta(G, X_0, X_n)] \geq -\mathbb{E}[H(X_n \mid (G, X_0))] \geq -\mathbb{E}[\log |B_{X_0}^G(n)|], \quad (3.7)$$

where  $H(\cdot \mid (G, X_0))$  denotes the Shannon entropy conditioned on  $(G, X_0)$ .

Using again the cycle property [BC12, Lem. 4.2], it holds that almost surely, for all  $n \geq 0$ ,

$$\log(\Delta(G, X_0, X_n)) = \sum_{t=0}^{n-1} \log(\Delta(G, X_t, X_{t+1})).$$

From (3.6) and the fact that  $\mathbb{E}[\log \deg_G(\rho)] < \infty$ , we have  $\mathbb{E}|\log \Delta(G, X_0, X_1)| < \infty$ . Thus using stationarity once more and combining this with (3.7) yields

$$\mathbb{E}[\log \Delta(G, X_0, X_1)] = \lim_{n \rightarrow \infty} \frac{\mathbb{E}[\log \Delta(G, X_0, X_n)]}{n} \geq \lim_{n \rightarrow \infty} \frac{-\mathbb{E}[\log |B_{X_0}^G(n)|]}{n} = 0,$$

implying that  $(G, X_0, X_1)$  is reversible (recall (3.4)).  $\square$

### 3.2 Choosing a good Følner set

Fix a rooted graph  $(G, \rho)$  with vertex set  $V_G$ . The most difficult part of the proof of Theorem 3.2 is verifying (1). Toward this end, we will employ Lemma 2.3 and Lemma 2.5 to control the random walk restricted to a subset of the vertices in  $G$  whenever the local growth rates are sufficiently well-behaved. Consider a finite subset  $S \subseteq V_G$ .

Let  $\{Z_t\}$  denote the stationary random walk restricted to  $S$  (recall Definition 1.12), and let  $\pi$  denote the corresponding stationary measure. Recall the definition of  $\phi_x^G(k)$  from (2.7). For  $k, k_0 \in \mathbb{N}$ , define the numbers

$$\begin{aligned} \bar{\phi}_S(k) &= \sum_{x \in S} \pi(x) \phi_x^G(k) \\ S_\lambda(k) &= \{x \in S : \phi_x^G(k) \leq \lambda \bar{\phi}_S(k)\} \\ S_\lambda^\uparrow(k_0) &= \bigcap_{k \geq k_0} S_{\lambda 2^{k-k_0}}(k). \end{aligned}$$

Note that by Markov's inequality and a geometric summation, we have

$$\pi(S_\lambda^\uparrow(k_0)) \geq 1 - 2/\lambda. \quad (3.8)$$

**Lemma 3.4.** *For all  $n \geq 1$ , it holds that*

$$\mathbb{E} \left[ d_G(Z_{2n}, Z_0)^2 \cdot \mathbf{1}_{S_\lambda^\uparrow(\alpha_n)}(Z_0) \right] \leq 2n + 256 \sum_{k=\alpha_n}^{\beta_n} 8^{2k} \exp \left( \frac{-8^{2k}}{\lambda^2 4^{k-\alpha_n+10} (1 + \bar{\phi}_S(k))^2 n} \right). \quad (3.9)$$

*Proof.* Apply Lemma 2.3 using the family of mappings that arises from applying Lemma 2.5 to  $G$ , and with the functional  $\varphi(k) = 128(1 + \lambda \bar{\phi}_S(k))$ .  $\square$

Therefore control on  $\bar{\phi}_S(k)$  for  $k \in \{\alpha_n, \dots, \beta_n\}$  yields control on the speed of  $\{Z_t\}$ . Let us define, for  $r \geq 1$ ,

$$\bar{\phi}_{\rho,r}^G(k) \stackrel{\text{def}}{=} \bar{\phi}_{B_\rho^G(r)}(k),$$

and chose  $S = B_\rho(r)$  for some  $r \geq 0$ .

**Definition 3.5** (Tempered growth). We say that a triple  $(n, \lambda, r)$  is *tempered* in  $(G, \rho)$  if

$$\bar{\phi}_{\rho,r}^G(k) \leq \lambda 2^{k-\alpha_n} \text{ for } k \in \{\alpha_n, \dots, \beta_n\}. \quad (3.10)$$

For ease of reference, we recall the definitions from (2.4)–(2.5):

$$\begin{aligned} \alpha_n &= \lceil \log_8(\sqrt{2n}) \rceil \\ \beta_n &= \lceil \log_8(2n) \rceil. \end{aligned}$$

**Lemma 3.6.** For all  $n \geq 1$  and  $\lambda \geq 2$  the following holds. If  $(n, \lambda, r)$  is tempered in  $(G, \rho)$  and  $S = B_\rho(r)$ , then

$$\mathbb{E} \left[ d_G(Z_{2n}, Z_0)^2 \cdot \mathbf{1}_{S_\lambda^\dagger(\alpha_n)}(Z_0) \right] \leq O(\lambda^{12})n.$$

*Proof.* To see this, apply Lemma 3.4 and note that  $8^{2\alpha_n} \leq O(n)$ , hence

$$\begin{aligned} \sum_{k=\alpha_n}^{\beta_n} 8^{2k} \exp \left( \frac{-8^{2k}}{\lambda^2 4^{k-\alpha_n+10} (1 + \bar{\phi}_S(k))^2 n} \right) &\leq O(n) \sum_{k=\alpha_n}^{\beta_n} 8^{2k-2\alpha_n} \exp \left( \frac{-8^{2k-2\alpha_n}}{\lambda^2 4^{k-\alpha_n+10} (1 + \lambda 2^{k-\alpha_n})^2} \right) \\ &\leq O(n) \sum_{j=0}^{\infty} 8^{2j} \exp \left( \frac{-4^j}{9 \cdot 4^{10} \lambda^4} \right) \\ &\leq O(\lambda^{12})n. \quad \square \end{aligned}$$

To compare the (unrestricted) random walk  $\{X_t\}$  on  $G$  to the walk  $\{Z_t\}$  restricted to  $B_\rho^G(r)$ , we will choose some  $r \geq 0$  satisfying

$$\frac{\mu_G(B_\rho^G(r) \setminus B_\rho^G(r-2n))}{\mu_G(B_\rho^G(r))} \leq \frac{1}{4\lambda}. \quad (3.11)$$

In particular, this implies that for  $\lambda \geq 1$ ,

$$\begin{aligned} \pi(B_\rho^G(r) \setminus B_\rho^G(r-2n)) &\leq \frac{\mu_G(B_\rho^G(r) \setminus B_\rho^G(r-2n))}{\mu_G(B_\rho^G(r)) - \mu_G(B_\rho^G(r) \setminus B_\rho^G(r-2n))} \\ &\stackrel{(3.11)}{\leq} \frac{1/4\lambda}{1 - 1/4\lambda} \\ &\leq \frac{1}{2\lambda}. \end{aligned} \quad (3.12)$$

**Definition 3.7** (Insulation). We say that a triple  $(n, \lambda, r)$  is *insulated* in  $(G, \rho)$  if (3.11) holds.

Our final choice of  $(n, \lambda, r)$  will satisfy some additional constraints, hence a complete description of the requirements is postponed to the next section. However we already have the following.

**Lemma 3.8.** For every  $\lambda \geq 2$  the following holds. For any rooted graph  $(G, \rho)$ , if  $(n, \lambda, r)$  is tempered and insulated in  $(G, \rho)$ , then for  $S = B_\rho^G(r)$ , it holds that

$$\pi \left( \left\{ x \in S : \mathbb{E}[d_G(X_{2n}, X_0)^2 \mid X_0 = x] \geq \lambda^{13} n \right\} \right) \leq O\left(\frac{1}{\lambda}\right).$$

*Proof.* Note that by [Lemma 3.6](#), and Markov's inequality,

$$\pi\left(\left\{x \in S_\lambda^\uparrow(\alpha_n) : \mathbb{E}[d_G(Z_{2n}, Z_0)^2 \mid X_0 = x] \geq \lambda^{13}n\right\}\right) \leq O\left(\frac{1}{\lambda}\right).$$

Now, unless  $x \in S \setminus B_\rho^G(r - 2n)$ , one can easily couple  $\{X_0, X_1, \dots, X_{2n}\}$  and  $\{Z_0, Z_1, \dots, Z_{2n}\}$  conditioned on  $X_0 = Z_0 = x$ . Thus using the fact that  $(n, \lambda, r)$  is tempered in  $(G, \rho)$ , along with [\(3.12\)](#) and [\(3.8\)](#), we have

$$\begin{aligned} \pi\left(\left\{x : \mathbb{E}[d_G(X_{2n}, X_0)^2 \mid X_0 = x] \geq \lambda^{13}n\right\}\right) &\leq \pi\left(S \setminus S_\lambda^\uparrow(\alpha_n)\right) + O\left(\frac{1}{\lambda}\right) + \pi(S \setminus B_\rho^G(r - 2n)), \\ &\leq O\left(\frac{1}{\lambda}\right). \end{aligned} \quad \square$$

Combining the preceding lemma with [\(3.11\)](#) gives us the following.

**Corollary 3.9.** *There is a constant  $\kappa > 0$  such that for every  $\lambda \geq 2$ , if  $(n, \lambda, r)$  is tempered and insulated in  $(G, \rho)$ , then*

$$\frac{\mu_G\left(\left\{x \in B_\rho^G(r) : \mathbb{E}[d_G(X_{2n}, X_0)^2 \mid X_0 = x] \geq \lambda^{13}n\right\}\right)}{\mu_G(B_\rho^G(r))} \leq \frac{\kappa}{\lambda}.$$

### 3.3 Multi-scale control of growth functionals

In this section we find scales which simultaneously satisfy all the criterion in [Theorem 3.2](#).

We begin with the following observation. For a function  $\phi : \mathbb{N} \rightarrow \mathbb{R}$ , an integer  $\ell \geq 1$ , let

$$\theta(\ell) = \sum_{k=\ell}^{3\ell} \phi(k) 2^{\ell-k}.$$

Then an elementary geometric summation yields

$$\sum_{\ell=h}^{2h} \theta(\ell) \leq 2 \sum_{k=h}^{5h} \phi(k). \quad (3.13)$$

Define now the quantity

$$\theta_{\rho,r}^G(\ell) = \sum_{k=\ell}^{3\ell} \bar{\phi}_{\rho,r}^G(k) 2^{\ell-k},$$

recalling that  $\bar{\phi}_{\rho,r}^G(k)$  is the average of  $\phi_x^G(k)$  over the stationary measure of the random walk restricted to  $B_\rho^G(r)$ . Note that

$$\theta_{\rho,r}^G(\ell) < \lambda \implies \bar{\phi}_{\rho,r}^G(k) < \lambda 2^{k-\ell} \text{ for } k \in \{\ell, \ell+1, \dots, 3\ell\}.$$

In particular, recalling [Definition 3.5](#),

$$\theta_{\rho,r}^G(\ell) < \lambda \implies (n, 128\lambda, r) \text{ is tempered in } (G, \rho) \text{ for } n \in [8^{2\ell}, 8^{2\ell+2}]. \quad (3.14)$$

Consider now a stationary random graph  $(G, \rho)$ . For  $k \geq 1$ , define

$$\hat{\psi}(k) = \mathbb{E}\left[\log |B_\rho^G(8^k)|\right].$$

If  $(G, \rho)$  has weakly annealed polynomial growth (1.4), then there is a number  $s > 0$  such that

$$\hat{\psi}(k) \leq sk. \quad (3.15)$$

We now fix a number  $k_0 \geq 3$ , and try to locate triples  $(n, \lambda, r)$  with  $n > 8^{k_0}$  that are tempered in  $(G, \rho)$  with high probability. To ensure simultaneous occurrence of the many conditions required, we define

$$\begin{aligned} \Psi_{(G, \rho)}(k_0) = \sum_{k=9k_0}^{10k_0} \left[ \log \frac{\mu_G(B_\rho^G(8^{k+2}))}{\mu_G(B_\rho^G(8^k))} + \left( \sum_{r \in I(k)} \log \frac{\mu_G(B_\rho^G(r))}{\mu_G(B_\rho^G(r - 8^{4k_0+3}))} \right. \right. \\ \left. \left. + \frac{1}{k_0 |I(k)|} \sum_{\ell=k_0}^{2k_0} \left[ \theta_{\rho, r}^G(\ell) + \left\{ \sum_{n=8^{2\ell}}^{8^{2\ell+2}} H_{(G, \rho)}(2n) - H_{(G, \rho)}(2n-1) \right\} \right] \right) \right], \end{aligned}$$

where  $I(k) = \{8^k + 8^{4k_0+3}, 8^k + 2 \cdot 8^{4k_0+3}, 8^k + 3 \cdot 8^{4k_0+3}, \dots, 8^{k+1} - 8^{4k_0+3}\}$ .

Observe that from (3.13), for any  $r \in [9k_0, 10k_0]$  we have,

$$\sum_{\ell=k_0}^{2k_0} \theta_{\rho, r}^G(\ell) \leq 2 \sum_{\ell=k_0}^{5k_0} \bar{\phi}_{B_\rho^G(r)}^G(\ell) \leq 2 \log |B_\rho^G(8^{5k_0} + r)| \leq 2 \log |B_\rho^G(8^{11k_0})|.$$

The sum in braces is bounded by  $H_{(G, \rho)}(2 \cdot 8^{4k_0+2})$  which is at most  $\log |B_\rho^G(2 \cdot 8^{4k_0+2})|$ . The first two terms sum telescopically to at most  $4 \log \mu_G(B_\rho^G(8^{10k_0+2}))$ . Putting everything together, we arrive at

$$\Psi_{(G, \rho)}(k_0) \leq 4 \log \mu_G(B_\rho^G(8^{10k_0+2})) + 3 \log |B_\rho^G(8^{11k_0})| \leq 11 \log |B_\rho^G(8^{11k_0})|. \quad (3.16)$$

Note that we use the trivial bound  $\mu_G(B_\rho^G(8^{10k_0+2})) \leq |B_\rho^G(8^{11k_0})|^2$  and the fact that  $k_0 \geq 3$ .

The growth assumption (1.4) now implies that

$$\gamma \stackrel{\text{def}}{=} \mathbb{E}[\Psi_{(G, \rho)}(k_0)] \leq 121k_0 s.$$

Thus there must exist numbers  $(k, r, \ell, n)$  with  $k \in [9k_0, 10k_0]$  and  $r \in [8^k, 8^{k+1}]$  such that

$$\mathbb{E} \left[ \log \frac{\mu_G(B_\rho^G(8^{k+2}))}{\mu_G(B_\rho^G(8^k))} \right] \leq \frac{4\gamma}{k_0} \leq O(s), \quad (3.17)$$

$$\mathbb{E} \left[ \log \frac{\mu_G(B_\rho^G(r))}{\mu_G(B_\rho^G(r - 8^{4k_0+3}))} \right] \leq \frac{4\gamma}{8^{5k_0-3}}, \quad (3.18)$$

and there are similarly  $\ell \in [k_0, 2k_0]$  and

$$n \in [8^{2\ell}, 8^{2\ell+2}] \quad (3.19)$$

such that

$$\mathbb{E} [\theta_{\rho, r}^G(\ell)] \leq \frac{4\gamma}{k_0} \leq O(s) \quad (3.20)$$

$$H_{2n} - H_{2n-1} = \mathbb{E} [H_{(G, \rho)}(2n) - H_{(G, \rho)}(2n-1)] \leq \frac{4\gamma}{k_0 8^{2\ell}} \leq O(s/n), \quad (3.21)$$

With the above preparation, we are now ready to finish the proof of Theorem 3.2.

*Proof of Theorem 3.2.* For every  $k_0 \geq 3$ , we obtain a quadruple  $(k(k_0), r(k_0), \ell(k_0), n(k_0))$  satisfying (3.17)–(3.18) and (3.20)–(3.21). Fix an infinite and strictly increasing sequence of values  $\{k_1, k_2, \dots\}$  so that the sequence of times

$$\mathbb{T} = \{2 \cdot n(k_i) : i = 1, 2, \dots\}$$

is also strictly increasing.

For  $2 \cdot n(k_i) \in \mathbb{T}$ , define  $r_{2n(k_i)} = r(k_i)$ . Let  $\{k_i, \ell_i, 2n_i\}$  denote the sequence  $\{k(k_i), \ell(k_i), 2n(k_i)\}$ . Inequalities (3.17) and (3.21) show that  $\mathbb{T}$  and  $\{r_n : n \in \mathbb{T}\}$  satisfy conditions (2) and (3) of Theorem 3.2 for some constant  $C > 0$ . It remains to verify condition (1).

Toward this end, consider some  $\varepsilon > 0$ . From (3.18) and (3.20), for every  $i = 1, 2, \dots$ , we can choose constants  $c \leq O(s/\varepsilon)$  and  $b_i \leq O(n_i^{-2/3}/\varepsilon)$  such that the event

$$\mathcal{E}_i = \left\{ \theta_{\rho, r_{n_i}}^G(\ell_i) < c \right\} \wedge \left\{ \log \frac{\mu_G(B_\rho^G(r_{n_i}))}{\mu_G(B_\rho^G(r_{n_i} - 8^{4k_i+3}))} < b_i \right\}$$

has  $\mathbb{P}(\mathcal{E}_i) \geq 1 - \frac{1}{2}\varepsilon$ .

Note that from (3.14) and the choice (3.19), we know that for  $i = 1, 2, \dots$ ,

$$\mathcal{E}_i \implies (n_i, 128c, r_{n_i}) \text{ is tempered in } (G, \rho).$$

Define  $\lambda = 2\kappa/\varepsilon$ , where  $\kappa$  is the constant from Corollary 3.9. Then since  $2n_i \leq 8^{2\ell_i+3} \leq 8^{4k_i+3}$ , it holds that

$$\begin{aligned} \log \frac{\mu_G(B_\rho^G(r_{n_i}))}{\mu_G(B_\rho^G(r_{n_i} - 8^{4k_i+3}))} < \frac{1}{4\lambda} &\implies \frac{\mu_G(B_\rho^G(r_{n_i}) \setminus B_\rho^G(r_{n_i} - 8^{4k_i+3}))}{\mu_G(B_\rho^G(r_{n_i}))} < \frac{1}{4\lambda} \\ &\implies \frac{\mu_G(B_\rho^G(r_{n_i}) \setminus B_\rho^G(r_{n_i} - 2n_i))}{\mu_G(B_\rho^G(r_{n_i}))} < \frac{1}{4\lambda}, \\ &\implies (n_i, \lambda, r_{n_i}) \text{ is insulated in } (G, \rho). \end{aligned}$$

where the first inequality uses  $\log(1-x) \leq -x$  for  $x \in [0, 1)$ . Therefore,

$$\mathcal{E}_i \wedge \left\{ b_i < \frac{1}{4\lambda} \right\} \implies (n_i, \lambda, r_{n_i}) \text{ is insulated in } (G, \rho).$$

Note that  $b_i < 1/(4\lambda)$  occurs for all  $k_i$  sufficiently large (since  $n_i \rightarrow \infty$  as  $k_i \rightarrow \infty$ ).

We can thus apply Corollary 3.9 to conclude that if  $\mathcal{E}_i$  occurs and  $k_i$  is sufficiently large, then there is a constant  $C' > 0$  such that

$$\frac{\mu_G(\{x \in B_\rho^G(r_{n_i}) : \mathbb{E}[d_G(X_{2n_i}, X_0)^2 \mid X_0 = x, (G, \rho)] \geq (C'\varepsilon^{-13})n_i\})}{\mu_G(B_\rho^G(r_{n_i}))} \leq \frac{\varepsilon}{2}.$$

We conclude that

$$\mathbb{E} \left[ \frac{\mu_G(\{x \in B_\rho^G(r_{n_i}) : \mathbb{E}[d_G(X_0, X_{2n_i})^2 \mid X_0 = x, (G, \rho)] \geq (C'\varepsilon^{-13})n_i\})}{\mu_G(B_\rho^G(r_{n_i}))} \right] \leq \frac{\varepsilon}{2} + (1 - \mathbb{P}[\mathcal{E}_i]) \leq \varepsilon,$$

completing the proof.  $\square$



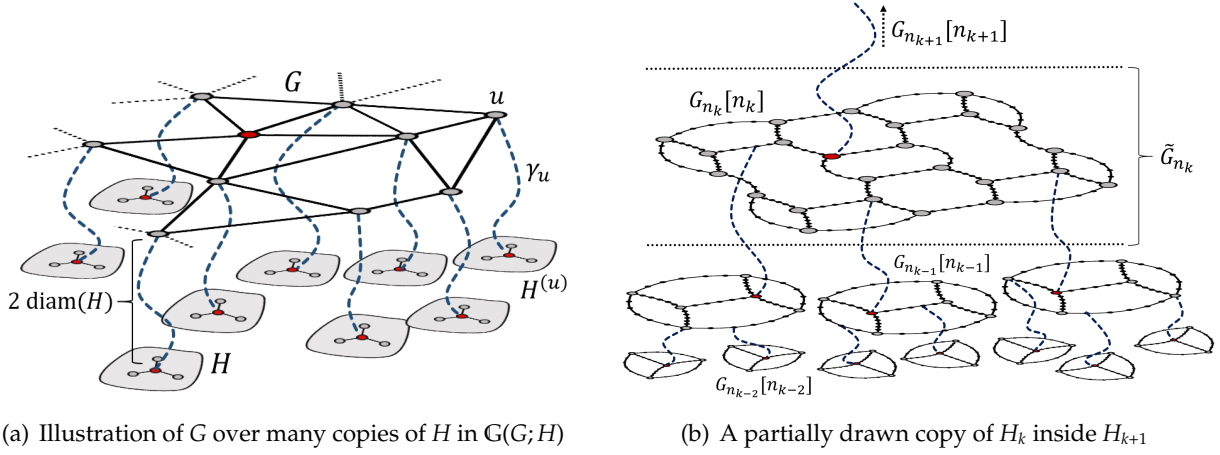


Figure 1: Trees of graphs

## 4 Existence of exceptional times

We now present an example showing that one cannot hope to prove [Theorem 1.6](#) for all times. For ease of notation, throughout this section, if  $G$  is a graph, we use  $V(G)$  and  $E(G)$  for the vertex and edge set of  $G$ , respectively.

**Theorem 4.1.** *There exists a stationary random rooted graph  $(G, \rho)$  with the following properties:*

1. *Almost surely: For any  $x \in V(G)$  and  $r > 0$ , it holds that  $|B_x^G(r)| \leq O(r^7)$ .*
2. *Almost surely:  $\sup_{x \in V(G)} \deg_G(x) = 4$ .*
3. *Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be an unbounded, monotone increasing function. Then there is a sequence of times  $\{t_k\}_{k \in \mathbb{N}}$  so that*

$$\lim_{k \rightarrow \infty} \mathbb{P} \left( \mathbb{E} \left[ d_G(X_0, X_{t_k})^2 \mid (G, \rho), X_0 = \rho \right] \geq t_k \frac{\log t_k}{f(t_k)} \right) = 1.$$

**Remark 4.2.** With more effort, it is possible to obtain a similar construction with  $|B_x^G(r)| \leq r^{2+o(1)}$ .

The basic idea of the construction is simple: Let  $G_n[n]$  denote the result of taking a 3-regular expander graph on  $n$  vertices and replacing every edge by a path of length  $n$ . Then by construction, the volume growth is at most quadratic, but after time  $\approx n^2 \log n$ , the random walk will have gone distance  $\approx n^2 (\log n)^2$ , making it slightly superdiffusive. The technical difficulties lie in converting this finite family of examples into a stationary random graph. To accomplish this, we build a tree of such graphs (see [Figure 1\(b\)](#)), with the sizes decreasing rapidly down the tree, and with buffers between the levels to enforce polynomial volume growth.

### 4.1 Trees of graphs

We first describe a certain way of constructing graphs from other graphs and provide some preliminary estimates on the properties of the construction. In this section, we will deal primarily with rooted graphs. For a graph  $G$ , we use  $\rho_G \in V(G)$  to denote its root.

**A tree of  $H$ 's under  $G$ .** Consider two rooted graphs  $H$  and  $G$ . Construct a new rooted graph  $\mathbb{G} = \mathbb{G}(G; H)$  as follows. Take  $|V(G)| - 1$  disjoint copies of  $H$ :  $\{H^{(u)} : u \in V(G) \setminus \{\rho_G\}\}$ . Let  $\rho_H^{(u)}$  be the copy of  $\rho_H$  in  $H^{(u)}$ . Let  $\{\gamma_u : u \in V(G) \setminus \{\rho_G\}\}$  be a collection of edge-disjoint paths of length  $2 \text{diam}(H)$  where  $\gamma_u$  connects  $u \in V(G) \setminus \{\rho_G\}$  to  $\rho_H^{(u)}$  in  $H^{(u)}$ . Define

$$\begin{aligned} V(\mathbb{G}) &= V(G) \cup \bigcup_{u \in V(G) \setminus \{\rho_G\}} \left( V(H^{(u)}) \cup V(\gamma_u) \right), \\ E(\mathbb{G}) &= E(G) \cup \bigcup_{u \in V(G) \setminus \{\rho_G\}} \left( E(H^{(u)}) \cup E(\gamma_u) \right). \end{aligned}$$

There is a natural identification  $V(G) \subseteq V(\mathbb{G})$  and we define the root  $\rho_{\mathbb{G}} = \rho_G$  of  $\mathbb{G}$ . We refer to the paths  $\{\gamma_u\}$  as *tails*. See [Figure 1\(a\)](#).

For a graph  $G$ , let us use  $\Delta_G = \max_{v \in V(G)} \deg_G(v)$  to denote its maximum degree.

**Lemma 4.3.** *For any rooted graphs  $H$  and  $G$  and  $h \geq 1$ , if  $\mathbb{G} = \mathbb{G}(G; H)$ , then  $\deg_{\mathbb{G}}(\rho_{\mathbb{G}}) = \deg_G(\rho_G)$ , and*

$$\begin{aligned} \Delta_{\mathbb{G}} &\leq \max\{\Delta_H, \Delta_G + 1, \deg_H(\rho_H) + 1\}. \\ |V(\mathbb{G})| &= |V(G)| + (|V(G)| - 1)(|V(H)| + 2 \text{diam}(H) - 2), \\ \text{diam}(\mathbb{G}) &\leq \text{diam}(G) + 6 \text{diam}(H). \end{aligned}$$

**Graph subdivision.** For a parameter  $L \in \mathbb{N}$ , we define a graph  $G[L]$  as the one which arises from  $G$  by subdividing every edge in  $E(G)$  into a path of length  $L$ . If  $G$  has root  $\rho_G$ , then under the natural identification  $V(G) \subseteq V(G[L])$ , we set  $\rho_{G[L]} = \rho_G$ . Note that:

$$|V(G[L])| = |V(G)| + |E(G)| \cdot (L - 1). \quad (4.1)$$

## 4.2 Stretched expanders and the rate of escape

Let  $\{G_n : n \in 2\mathbb{N}\}$  denote a family of 3-regular,  $n$ -vertex non-bipartite expander graphs. For each such  $n$ , we distinguish an arbitrary root  $\rho_{G_n} \in V(G_n)$ . We use  $t_{\text{mix}}(G)$  to denote the (total variation) mixing time of a graph  $G$ .

**Fact 4.4.** *There is a constant  $C > 0$  such that  $t_{\text{mix}}(G_n) \leq C \log n$  for all  $n \in 2\mathbb{N}$ .*

Since  $G_n$  is 3-regular, a fixed vertex is further than  $\frac{1}{3} \log n$  from all but  $o(n)$  vertices in  $G_n$ . Combining this with the preceding fact yields the following.

**Lemma 4.5.** *There is a constant  $c \geq 1$  such that the following holds for every  $\varepsilon \in (0, 1)$  and  $n \geq 1/\varepsilon^3$ . If  $\{X_t\}$  is the random walk in  $G_n$  and  $t \geq (c/\varepsilon)t_{\text{mix}}(G_n)$ , then*

$$\min_{x \in V(G_n)} \mathbb{P} \left[ d_{G_n}(X_0, X_t) > \frac{1}{3} \log n \mid X_0 = x \right] \geq 1 - \varepsilon. \quad (4.2)$$

For a graph  $H$ , define

$$t_{\text{hit}}(H) = \max_{x, y \in V(H)} \mathbb{E} \left[ \min \{t \geq 0 : Y_t = y\} \mid Y_0 = x \right],$$

where  $\{Y_t\}$  is the random walk on  $H$ .

One has the following basic estimate (see, e.g., [\[LPW09, Ch. 10\]](#)):

$$t_{\text{hit}}(H) \leq 2\Delta_H |V(H)|^2. \quad (4.3)$$

Let  $\pi_G$  denote the stationary measure of the random walk on a graph  $G$ .

**Lemma 4.6.** *There is a constant  $c' > 0$  such that the following holds for all  $\varepsilon \in (0, 1)$  and  $n \geq 1/\varepsilon^3$ . Let  $\{X_t\}$  denote the random walk on  $\mathbb{G} = \mathbb{G}(G_n[L]; H)$ , with  $X_0$  chosen according to the stationary measure  $\pi_{\mathbb{G}}$ . Assume that  $L \geq \text{diam}(H)$  and  $\log n \geq 48$ . Then for all*

$$t \in \left[ \frac{c'}{\varepsilon^2} \Delta_H |V(H)|^2 L^2 t_{\text{mix}}(G_n), \frac{\varepsilon L^2 n}{c'} \right],$$

it holds that

$$\mathbb{P} \left( \mathbb{E} \left[ d_{\mathbb{G}}(X_t, X_0)^2 \mathbf{1}_{\{\rho_{\mathbb{G}} \notin \{X_0, X_1, \dots, X_t\}\}} \mid X_0 \right] \geq \frac{(L \log n)^2}{72} \right) \geq 1 - \varepsilon.$$

*Proof.* Let  $t_{\text{mix}} = (3c/\varepsilon)t_{\text{mix}}(G_n)$ . Let  $\hat{\mathbb{G}}$  denote the graph  $\mathbb{G}$ , but where a path  $\hat{\gamma}$  of length  $2 \text{diam}(H)$  is added between  $\rho_{\mathbb{G}}$  and a new copy  $H^{(\rho_{\mathbb{G}})}$  of  $H$  (so that now all vertices of  $G_n[L]$  have a copy of  $H$  attached).

Consider the random walk  $\{Z_k\}$  on  $\hat{\mathbb{G}}$ . Let  $\tau_1 < \tau_2 < \dots$  be the sequence of times at which  $Z_{\tau_j} \in V(G_n)$ . Let  $K = \max\{j : \tau_j < t\}$  (and let  $K = 0$  if no such  $j$  exists). Observe that, conditioned on the sequence  $\{\tau_j\}$ , the process  $\{Z_{\tau_1}, Z_{\tau_2}, \dots\}$  has the law of random walk on  $G_n$ , therefore using (4.2) yields

$$\mathbb{P} \left[ \{K > 0\} \wedge d_{G_n}(Z_{\tau_K}, Z_{\tau_1}) \geq \frac{1}{3} \log n \mid Z_0 \right] \geq (1 - \frac{1}{3}\varepsilon) \mathbb{P}[K \geq t_{\text{mix}} \mid Z_0]. \quad (4.4)$$

The waiting periods  $\{\tau_{j+1} - \tau_j : j = 1, 2, \dots\}$  are i.i.d., and we have the estimates

$$\mathbb{E}[\tau_1 \mid Z_0], \mathbb{E}[\tau_{j+1} - \tau_j] \stackrel{(4.3)}{\leq} 3\Delta_H L^2 |V(H)|^2.$$

Hence,

$$\mathbb{P}[K < t_{\text{mix}} \mid Z_0] \leq \mathbb{P}[\tau_{\lfloor t_{\text{mix}} \rfloor} > t \mid Z_0] \leq \frac{\mathbb{E}[\tau_{\lfloor t_{\text{mix}} \rfloor} \mid Z_0]}{t} \leq \frac{3\Delta_H L^2 |V(H)|^2 t_{\text{mix}}}{t}.$$

Combined with (4.4), this shows for  $t \geq (9/\varepsilon)\Delta_H L^2 |V(H)|^2 t_{\text{mix}}$ ,

$$\mathbb{P} \left[ \{K > 0\} \wedge d_{G_n}(Z_{\tau_1}, Z_{\tau_K}) \geq \frac{1}{3} \log n \mid Z_0 \right] \geq (1 - \frac{1}{3}\varepsilon)^2 \geq 1 - \frac{2}{3}\varepsilon. \quad (4.5)$$

Now, note that as long as  $K > 0$  and  $\rho_{\mathbb{G}} \notin \{Z_0, \dots, Z_t\}$ , we can couple  $\{Z_0, \dots, Z_t\}$  with the random walk  $\{X_0, \dots, X_t\}$  on  $\mathbb{G}$ . For any  $T \geq 1$ ,

$$\begin{aligned} \mathbb{P}[\rho_{\mathbb{G}} \in \{Z_0, Z_1, \dots, Z_t\}] &\leq \mathbb{P}[K > T] + \mathbb{P}[\rho_{\mathbb{G}} \in \{Z_{\tau_1}, Z_{\tau_2}, \dots, Z_{\tau_K}\} \mid K \leq T] \\ &\leq \mathbb{P}[K > T] + \frac{(T+1)}{n}, \end{aligned} \quad (4.6)$$

where the last inequality follows because  $\{Z_{\tau_1}, \dots, Z_{\tau_K}\}$  is a stationary walk on  $G_n$ , conditioned on  $\{\tau_j\}$ , and because  $G_n$  is regular.

We now require a basic estimate on  $\tau_{j+1} - \tau_j$ . Let  $Y$  denote the amount of time needed for a random walk on  $\mathbb{Z}$ , started at the origin, to hit the set  $\{-L, L\}$ . Then  $\tau_{j+1} - \tau_j$  stochastically dominates  $Y$ , and we have the standard identities (see, e.g., [Moo73]):

$$\begin{aligned} \mathbb{E}[Y] &= L^2 \\ \text{Var}(Y) &= \frac{2(L^4 - L^2)}{3}. \end{aligned}$$

Let  $\{Y_j\}$  be i.i.d. copies of  $Y$ , and use Chebyshev's inequality to obtain:

$$\mathbb{P} \left( \tau_{m+1} < \frac{1}{2} mL^2 \right) \leq \mathbb{P} \left( Y_1 + \dots + Y_m < m \mathbb{E}[Y] - \frac{1}{2} mL^2 \right) \leq \frac{8}{3m}.$$

For  $T = \lceil \frac{2t}{L^2} \rceil$ , this yields

$$\mathbb{P}(K > T) \leq \mathbb{P}(\tau_{T+1} < t) \leq \frac{8}{3T} \leq \frac{2L^2}{t}.$$

Plugging this into (4.6) gives

$$p := \mathbb{P}[\rho_G \in \{Z_0, Z_1, \dots, Z_t\}] \leq \frac{2L^2}{t} + \frac{t}{L^2 n}.$$

Note that  $p \leq \varepsilon/6$  as long as  $t \in [\frac{24L^2}{\varepsilon}, \frac{\varepsilon L^2 n}{12}]$ .

Using this in conjunction with (4.5), we arrive at

$$\begin{aligned} \mathbb{P}\left(\mathbb{P}\left[\{K > 0\} \wedge \rho_G \notin \{X_0, X_1, \dots, X_t\} \wedge d_{G_n}(X_{\tau_1}, X_{\tau_K}) \geq \frac{\log n}{3} \mid X_0\right] \geq \frac{1}{2}\right) &\geq (1 - \frac{2}{3}\varepsilon) - 2p \\ &\geq 1 - \varepsilon. \end{aligned} \quad (4.7)$$

When  $K > 0$ , the triangle inequality gives us

$$d_G(X_0, X_t) \geq L \cdot d_{G_n}(X_{\tau_1}, X_{\tau_K}) - 2(L + 3 \operatorname{diam}(H)).$$

Combining this with (4.7), along with the assumptions that  $L \geq \operatorname{diam}(H)$  and  $\log n \geq 48$  yields

$$\mathbb{P}\left(\mathbb{E}\left[d_{G_n}(X_0, X_t)^2 \mathbf{1}_{\{\rho_G \notin \{X_0, X_1, \dots, X_t\}\}} \mid X_0\right] \geq \frac{1}{2} \frac{(L \log n)^2}{36}\right) \geq 1 - \varepsilon,$$

completing the proof.  $\square$

### 4.3 The recursive construction

Observe that there is a constant  $C > 0$  such that for  $n \geq 4$ ,

$$\operatorname{diam}(G_n) \leq C \log n. \quad (4.8)$$

Let us denote  $n_0 = 10$  and suppose that  $n_k \geq 2n_{k-1}^2$  for  $k \geq 1$ . We define an inductive sequence of rooted graphs  $\{H_k\}$  as follows:  $H_0$  is the graph consisting of a single vertex, and for  $k \geq 1$ ,

$$H_k = G(G_{n_k}[n_k]; H_{k-1}).$$

Refer to [Figure 1\(b\)](#) for a depiction.

We begin by consulting [Lemma 4.3](#) for the following estimates. Using (4.8), we have for  $k \geq 1$ :

$$\operatorname{diam}(H_k) \leq \operatorname{diam}(G_{n_k}[n_k]) + 6 \operatorname{diam}(H_{k-1}) \leq C n_k \log n_k + 6 \operatorname{diam}(H_{k-1}).$$

Thus one can easily verify by induction that

$$\operatorname{diam}(H_k) \leq 2C n_k \log n_k. \quad (4.9)$$

Moreover,

$$|V(H_k)| \leq 2n_k^2 (|V(H_{k-1})| + 2 \operatorname{diam}(H_{k-1})) \leq 2n_k^2 (|V(H_{k-1})| + 4C n_{k-1} \log n_{k-1}),$$

and one verifies by induction that for  $k \geq 1$ ,

$$n_k^2 \leq |V(H_k)| \leq 2C n_k^4. \quad (4.10)$$

From [Lemma 4.3](#), the following bound on the vertex degrees is immediate:

$$\sup_{k \geq 1} \Delta_{H_k} \leq 4. \quad (4.11)$$

**Levels of vertices.** The graph  $H_k$  consists of a copy of  $G_{n_k}[n_k]$  connected to  $|V(G_{n_k}[n_k])| - 1$  copies of  $H_{k-1}$  via tails of length  $2 \operatorname{diam}(H_{k-1})$ . Each such copy of  $H_{k-1}$  contains a copy of  $G_{n_{k-1}}[n_{k-1}]$  that is connected to  $|V(G_{n_{k-1}}[n_{k-1}])| - 1$  copies of  $H_{k-2}$  via tails of length  $2 \operatorname{diam}(H_{k-2})$ , and so on. If a tail connects  $G_{n_\ell}[n_\ell]$  to  $H_{\ell-1}$ , we refer to it as a *level- $\ell$  tail*.

Naturally, we can think of every vertex as occurring in either in a copy of  $G_{n_\ell}$  for some  $\ell \leq k$ , or in a tail between  $G_{n_\ell}[n_\ell]$  and a copy of  $H_{\ell-1}$ . Let  $\mathcal{T}_\ell^k \subseteq V(H_k)$  denote the set of internal vertices in level- $\ell$  tails. Let  $V_\ell^k \subseteq V(H_k)$  denote the set of vertices occurring in some copy of  $G_{n_\ell}[n_\ell]$ . Note that the sets  $\{V_\ell^k, \mathcal{T}_\ell^k : \ell = 1, 2, \dots, k\}$  form a partition of  $V(H_k)$ .

Finally, we use the notation  $\tilde{G}_{n_\ell}$  for the graph  $G_{n_\ell}[n_\ell]$ , together with the tail of length  $\operatorname{diam}(H_\ell) - 1$  attached to  $\rho_{G_{n_\ell}[n_\ell]}$ , and the disjoint tails of length  $\operatorname{diam}(H_{\ell-1}) + 1$  attached to all vertices of  $V(G_{n_\ell}[n_\ell]) \setminus \{\rho_{G_{n_\ell}[n_\ell]}\}$ . Observe that  $V(H_k)$  partitions into a *disjoint union* of copies of  $\tilde{G}_{n_\ell}$  with  $\ell \in \{1, 2, \dots, k\}$ . Accordingly, we can write  $\lambda(x)$  for the index  $\ell$  such that  $x$  is in a copy of  $\tilde{G}_{n_\ell}$ , and  $\tilde{G}_x$  for the subgraph corresponding to  $x$ 's copy of  $\tilde{G}_{n_{\lambda(x)}}$ . We now observe the main point of the tails.

**Lemma 4.7.** *Suppose that  $x, y \in V(H_k)$ . If  $\tilde{G}_x \neq \tilde{G}_y$ , then either  $x$  and  $y$  lie on a common tail, or  $d_{H_k}(x, y) \geq \operatorname{diam}(H_{\min(\lambda(x), \lambda(y))})$ .*

**Lemma 4.8.** *For any  $n \in 2\mathbb{N}$ ,  $v \in V(\tilde{G}_n)$ , and  $r \geq 1$ , it holds that*

$$|B_v^{\tilde{G}_n}(r)| \leq 3r^3.$$

*Proof.* If  $r \leq n$ , then

$$|B_v^{\tilde{G}_n}(r)| \leq r |B_v^{G_n[n]}(r)| \leq r^2 \Delta_{G_n} \leq 3r^2.$$

Otherwise,  $|B_v^{\tilde{G}_n}(r)| \leq r |V(G_n[n])| \leq 2rn^2 \leq 2r^3$ .  $\square$

Observe also the basic estimate: For  $k \geq 1$ ,

$$\operatorname{diam}(H_k) \geq n_k \operatorname{diam}(G_{n_k}) \geq 2n_k. \quad (4.12)$$

**Lemma 4.9** (Polynomial volume growth). *For every  $k \geq 1$ ,  $r \geq 0$ , and  $x \in V(H_k)$ , it holds that*

$$|B_x^{H_k}(r)| \leq 6Cr^7.$$

*Proof.* Consider  $n_{\ell-1} \leq r < n_\ell$ . The main idea is that if  $j > \ell$ , then from [Lemma 4.7](#), we know the ball  $B_x^{H_k}(r)$  cannot intersect both the top and bottom half of a level- $j$  tail  $\gamma$  unless  $B_x^{H_k}(r) \subseteq V(\gamma)$ , because the length of  $\gamma$  is at least  $2 \operatorname{diam}(H_\ell) \geq 4n_\ell$ .

More precisely, from [Lemma 4.7](#) we know that one of the following cases occurs:

1. The ball  $B_x^{H_k}(r)$  is completely contained in some tail. In this case, clearly  $|B_x^{H_k}(r)| \leq 2r$ .
2. The ball  $B_x^{H_k}(r)$  is contained in a copy of  $\tilde{G}_{n_j}$  for an index  $j > \ell$ . In this case,  $|B_x^{H_k}(r)| \leq 3r^3$  from [Lemma 4.8](#).
3. It holds that  $\max \{ \lambda(v) : v \in B_x^{H_k}(r) \} \leq \ell$ . In this case, use [Lemma 4.8](#) to write

$$|B_x^{H_k}(r)| \leq \left( \max_{v \in V(\tilde{G}_{n_\ell})} |B_v^{\tilde{G}_{n_\ell}}(r)| \right) |V(H_{\ell-1})| \leq 3r^3 |V(H_{\ell-1})| \stackrel{(4.10)}{\leq} 6Cr^3 n_{\ell-1}^4 \leq 6Cr^7. \quad \square$$

Let  $\{X_t\}$  denote the random walk on  $H_k$  where  $X_0$  has law  $\pi_{H_k}$ .

**Lemma 4.10** (Speed of the random walk). *There is a constant  $c > 0$  such that the following holds: For all  $\varepsilon > 0$  and  $\ell$  sufficiently large (with respect to  $\varepsilon$ ), if  $k \geq \ell$  and*

$$t \in \left[ \frac{c}{\varepsilon^2} n_\ell^2 (\log n_\ell) |V(H_{\ell-1})|^2, \frac{\varepsilon n_\ell^3}{c} \right],$$

then

$$\mathbb{P} \left( \mathbb{E} [d_{H_k}(X_0, X_t)^2 \mid X_0] \geq \frac{(n_\ell \log n_\ell)^2}{72} \right) \geq 1 - \varepsilon.$$

*Proof.* For  $1 \leq \ell \leq k$ , let  $p_{\ell,k}$  denote the probability that a vertex  $v \in V(H_k)$  chosen uniformly at random does not fall in some copy of  $H_\ell$ . First, we use (4.10) and (4.9) to bound

$$p_{k-1,k} \leq \frac{|V(G_{n_k}[n_k])| \cdot 2 \operatorname{diam}(H_{k-1})}{(|V(G_{n_k}[n_k])| - 1) |V(H_{k-1})|} \leq \frac{O(\log n_{k-1})}{n_{k-1}}, \quad (4.13)$$

This yields

$$p_{\ell,k} \leq \sum_{j=\ell}^{k-1} p_{j,j+1} \leq O(1) \sum_{j=\ell}^{k+1} \frac{\log n_j}{n_j}.$$

Observe that since  $\Delta_{H_k} \leq 4$ , the probability that a vertex chosen from the stationary measure does not fall in some copy of  $H_\ell$  is bounded by  $4p_{\ell,k}$ . Recall that the sequence  $\{n_j\}$  is increasing rapidly:  $n_{j+1} \geq 2n_j^2$ . Let  $\ell$  be chosen large enough so that  $4p_{\ell,k} < \varepsilon$ .

Let  $\mathcal{E}(\ell)$  denote the event that  $X_0$  lies in a copy of  $H_\ell^*$  of  $H_\ell$ . We have  $\mathbb{P}[\mathcal{E}(\ell)] \geq 1 - \varepsilon$ . Moreover, conditioned on  $\mathcal{E}(\ell)$ , if the random walk  $\{X_0, X_1, \dots, X_t\}$  avoids the root  $\rho_{H_\ell^*}$ , then it can be coupled to a stationary random walk on  $H_\ell$ . Now applying Lemma 4.6 with  $G = H_\ell$  yields the desired result.  $\square$

#### 4.4 Convergence to a stationary random graph

Let  $\rho_k \in V(H_k)$  be chosen according to the stationary measure  $\pi_{H_k}$ , and let  $\mu_k$  be the law of the random rooted graph  $(H_k, \rho_k)$ .

**Lemma 4.11.** *The measure  $\mu := \lim_{k \rightarrow \infty} \mu_k$  exists in the local weak topology. Moreover, if  $(G, \rho)$  has the law of  $\mu$ , then  $(G, \rho)$  is a stationary random graph such that, almost surely,  $\sup_{x \in V(G)} \deg_G(x) \leq 4$ , and  $|B_\rho^G(r)| \leq 6Cr^7$  for all  $r \geq 1$ .*

*Proof.* Assuming that the limit exists, the latter assertions follow from (4.11) and Lemma 4.9.

By definition of the local weak topology, to prove convergence of the measures  $\mu_k$ , it suffices to show that for every  $r > 0$ , the measures  $\mu_{k,r}$  converge, where  $\mu_{k,r}$  is the law of  $B_{\rho_k}^{H_k}(r)$ . A standard application of Kolmogorov's extension theorem then proves the existence of the limit  $\mu$ . For more details, see [BS01].

Let  $\mathcal{E}$  denote the event that  $\rho_k$  lies in a copy of  $H_{k-1}$ . Observe that (recall (4.13)):

$$\mathbb{P}[\neg \mathcal{E}] \leq \frac{\Delta_{H_k} |V(G_{n_k}[n_k])| \cdot 2 \operatorname{diam}(H_{k-1})}{(|V(G_{n_k}[n_k])| - 1) |V(H_{k-1})|} \leq O\left(\frac{\log n_{k-1}}{n_{k-1}}\right).$$

Suppose that  $\mathcal{E}$  occurs, and let  $H_{k-1}^*$  denote the copy of  $H_{k-1}$  in  $H_k$  containing  $\rho_k$ . In this case, we can couple  $\rho_k$  and  $\rho_{k-1} \in V(H_{k-1})$  in the obvious way. Note furthermore that  $(B_{\rho_k}^{H_k}(r), \rho_k)$  and



$(B_{\rho_{k-1}}^{H_{k-1}}(r), \rho_{k-1})$  are coupled (under the natural isomorphism) as long as  $d_{H_{k-1}}(\rho_{k-1}, \rho_{H_{k-1}}) > r$ . This yields

$$\mathbb{P} \left[ d_{H_{k-1}}(\rho_{k-1}, \rho_{H_{k-1}}) \leq r \right] \leq \frac{|B_{\rho_{H_{k-1}}}^{H_{k-1}}(r)|}{|V(H_{k-1})|} \leq \frac{4^r}{n_{k-1}^2},$$

since  $\Delta_{H_{k-1}} \leq 4$ . We conclude that, for fixed  $r > 0$ , it holds that

$$d_{TV}(\mu_{k-1,r}, \mu_{k,r}) \leq O\left(\frac{\log n_{k-1}}{n_{k-1}}\right).$$

Since  $\frac{\log n_k}{n_k}$  is summable, this yields the desired convergence as  $k \rightarrow \infty$ .  $\square$

We are ready to complete the proof of [Theorem 4.1](#).

*Proof of Theorem 4.1.* Let  $(G, \rho)$  be the limit of  $(H_k, \rho_k)$  constructed in [Lemma 4.11](#). Properties (1) and (2) are satisfied by the statement of the lemma. Let  $\{\varepsilon_\ell\}$  denote a sequence with  $\varepsilon_\ell \rightarrow 0$  as  $\ell \rightarrow \infty$  and such that [Lemma 4.10](#) applies to  $\varepsilon = \varepsilon_\ell$  for  $k \geq \ell$ .

The third property follows from [Lemma 4.10](#) by choosing the sequence  $\{n_\ell\}$  to grow fast enough so that

$$|V(H_{\ell-1})|^2 \stackrel{(4.10)}{\leq} O(n_{\ell-1}^8) \leq \frac{o(\log n_\ell)}{\varepsilon_\ell^2 f(n_\ell)}.$$

as  $\ell \rightarrow \infty$ .  $\square$

## References

- [AL07] David Aldous and Russell Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007.
- [Bal92] K. Ball. Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2(2):137–172, 1992.
- [BC12] Itai Benjamini and Nicolas Curien. Ergodic theory on stationary random graphs. *Electron. J. Probab.*, 17:no. 93, 20, 2012.
- [BDCKY15] Itai Benjamini, Hugo Duminil-Copin, Gady Kozma, and Ariel Yadin. Disorder, entropy and harmonic functions. *Ann. Probab.*, 43(5):2332–2373, 2015.
- [Bow95] B. H. Bowditch. A short proof that a subquadratic isoperimetric inequality implies a linear one. *Michigan Math. J.*, 42(1):103–107, 1995.
- [BP89] Martin T. Barlow and Edwin A. Perkins. Symmetric Markov chains in  $\mathbb{Z}^d$ : how fast can they move? *Probab. Theory Related Fields*, 82(1):95–108, 1989.
- [BS01] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.
- [Car85] Thomas Keith Carne. A transmutation formula for Markov chains. *Bull. Sci. Math. (2)*, 109(4):399–405, 1985.
- [CKR01] Gruia Calinescu, Howard Karloff, and Yuval Rabani. Approximation algorithms for the 0-extension problem. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 8–16, Philadelphia, PA, 2001.

- [DLP13] Jian Ding, James R. Lee, and Yuval Peres. Markov type and threshold embeddings. *Geom. Funct. Anal.*, 23(4):1207–1229, 2013.
- [HSC93] W. Hebisch and L. Saloff-Coste. Gaussian estimates for Markov chains and random walks on groups. *Ann. Probab.*, 21(2):673–709, 1993.
- [Kes59] Harry Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92:336–354, 1959.
- [Kes86] Harry Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487, 1986.
- [Kle10] Bruce Kleiner. A new proof of Gromov’s theorem on groups of polynomial growth. *J. Amer. Math. Soc.*, 23(3):815–829, 2010.
- [KLMN05] R. Krauthgamer, J. R. Lee, M. Mendel, and A. Naor. Measured descent: A new embedding method for finite metrics. *Geom. Funct. Anal.*, 15(4):839–858, 2005.
- [LP16] Russell Lyons and Yuval Peres. *Probability on Trees and Networks*. Cambridge University Press, 2016. Available at <http://pages.iu.edu/~rdlyons/>.
- [LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
- [LZ94] T. J. Lyons and T. S. Zhang. Decomposition of Dirichlet processes and its application. *Ann. Probab.*, 22(1):494–524, 1994.
- [Moo73] J. W. Moon. Random walks on random trees. *J. Austral. Math. Soc.*, 15:42–53, 1973.
- [NPSS06] Assaf Naor, Yuval Peres, Oded Schramm, and Scott Sheffield. Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.*, 134(1):165–197, 2006.
- [Pin94] Iosif Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.*, 22(4):1679–1706, 1994.
- [ST10] Yehuda Shalom and Terence Tao. A finitary version of Gromov’s polynomial growth theorem. *Geom. Funct. Anal.*, 20(6):1502–1547, 2010.
- [Var85] Nicholas Th. Varopoulos. Long range estimates for Markov chains. *Bull. Sci. Math. (2)*, 109(3):225–252, 1985.
- [Woe00] Wolfgang Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2000.